Symbolische Repräsentationen von Grauwertbildern für Szenenanalysen

von

Jörgen Peter Foith

Erlangen, 1978

1978 № 2 67

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41548222_0081

Symbolische Repräsentationen von Grauwertbildern für Szenenanalysen

Dem Fachbereich Ingenieurwissenschaften
der Universität Erlangen-Nürnberg
zur Erlangung des Grades
Doktor - Ingenieur

vorgelegt
von

Jörgen Peter Foith

Erlangen, 1978

UNIV.
LEIDEN
BIBL.

Als Dissertation genehmigt vom
Fachbereich Ingenieurwissenschaften
der Universität Erlangen-Nürnberg

Tag der Einreichung : 6. Januar 1978
Tag der Promotion : 13. März 1978
Dekan : Prof. Dr. D. Seitzer
Berichterstatter : Prof. Dr. H. Niemann
Prof. Dr. H.J. Schneider

INHALTSVERZEICHNIS SEITE

1. EINLEITUNG 1

1.1 Perzeptive und kognitive Prozesse beim Menschen und in der Technik 1

1.2 Zielsetzung und Beitrag der Arbeit 3

2. SZENENANALYSE BEIM MENSCHEN 5

3. SZENENANALYSE IM COMPUTER 9

3.1 Verfahren der Literatur 9

3.2 Experimentelle Erprobung und Kritik ausgewählter Verfahren 15

3.3 Ein Katalog von Verarbeitungsprinzipien 28

4. RELATIONAL-BÄUME: EIN NEUER ANSATZ ZUR REPRÄSENTATION VON GRAUWERTBILDERN 34

4.1 Die Repräsentation von Bildzeilen durch Relational-Bäume 36

4.2 Verallgemeinerte Relational-Bäume 43

4.3 Der Algorithmus zur Konstruktion von Relational-Bäumen 47

5. DIE AUSWERTUNG VON RELATIONAL-BÄUMEN 61

5.1 Die Rekonstruktion von Intensitätsprofilen aus Relational-Bäumen 62

5.2 Kantendetektion 66

5.3 Gruppierungs-Mechanismen 72

5.3.1 Gruppierungen 72

5.3.2 Tiefenstrukturen 82

INHALTSVERZEICHNIS SEITE

5.4 Texturanalysen 85

5.5 Die Verarbeitung quer zu den Zeilen: Profilverfolgung 101

6. EIN MODELL NIEDERER PROZESSE FÜR SZENENANALYSEN 107

6.1 Die Kontroll-Struktur des Modells 107

6.2 Die Skizze als Sekundär-Repräsentation und ihre Verarbeitung 112

6.3 Diskussion 116

7. ZUSAMMENFASSUNG 120

8. LITERATURVERZEICHNIS 123

SCHLUSSBEMERKUNG 131

LEBENSLAUF 132

1. EINLEITUNG

Der Mensch ist in der Lage, Szenen - d.h. Bilder, die eine Vielzahl verschiedener Objekte enthalten - zu analysieren und zu beschreiben. Eine automatische Analyse von Szenen mit dem Computer wird wegen der Vielfalt möglicher Anwendungen angestrebt. Es ist Aufgabe der Szenenanalyse, eine Szene in ihre Bestandteile zu zerlegen und eine Beschreibung über Art, Anzahl, Lage und Zustand der in der Szene enthaltenen Objekte zu liefern. Die Verarbeitungsschritte werden in zwei Gruppen eingeteilt: (1) in "niedere Prozesse", die eine Szene bis zur Ebene der Segmentation und Merkmalsextraktion verarbeiten, und (2) in "höhere Prozesse", die sich mit der Zuordnung segmentierter Regionen zu internen Modellen und dem Aufbau interner Modelle befassen. Im Arbeitsgebiet der Künstlichen Intelligenz wurde eine Vielzahl von Verfahren für Szenenanalysen entwickelt - vornehmlich in den USA. In der noch kurzen Historie der Entwicklung zeichnen sich drei Phasen ab: in der ersten Phase wurden einfache künstliche Szenen durch Algorithmen verarbeitet, die durch lokale Operationen und Entscheidungen gekennzeichnet waren. Eine solche Vorgehensweise erwies sich sogar für diese Szenen als nur bedingt brauchbar. Durch Ausnützen von A-priori-Wissen und die Beeinflussung niederer Prozesse durch interne Modelle wurde in der zweiten Phase eine Verbesserung erzielt. In den letzten Jahren verlagerte sich das Interesse von der Analyse künstlicher Szenen zur Auswertung von natürlichen und komplexen Szenen. Hier ist der Einsatz von A-priori-Wissen und internen Modellen für die Steuerung niederer Prozesse nur in Spezialfällen möglich. Deshalb wird in der dritten Phase versucht, autonome niedere Prozesse zu entwickeln, die ihre Entscheidungen mit Hilfe globaler Information treffen. Die Extraktion globaler Information wird durch hierarchische Repräsentationen der natürlichen Szenen ermöglicht. Die meisten der bisher entwickelten Repräsentationen bauen auf Zerlegungen von Szenen aufgrund einer fest vorgegebenen Geometrie auf. Die vorliegende Arbeit befaßt sich mit hierarchischen Repräsentationen und ihrer Auswertung. Im Gegensatz zu vielen bekannten Ansätzen werden hier kontext-abhängige Verfahren entwickelt.

1.1 PERZEPTIVE UND KOGNITIVE PROZESSE BEIM MENSCHEN UND IN DER TECHNIK

Im folgenden wird der Sprachgebrauch für einige wichtige Begriffe festgelegt. Eine Szene ist die zweidimensionale Darstellung eines Ausschnittes der dreidimensionalen, optisch erfaßbaren Umwelt. Für die Lösung technischer Aufgaben ist es notwendig, Art und Anzahl der Objekte einer Szene sowie die Komplexität ihrer Konfiguration stark einzuschränken.

Beim Menschen werden perzeptive und kognitive Prozesse unterschieden, wobei eine strenge Abgrenzung zwischen beiden Prozessen beim heutigen Stand des Wissens nicht gegeben werden kann. Perzeptive Prozesse sind Vorgänge, die ohne interne Modelle auskommen; kognitive Prozesse operieren mit internen Modellen. Unter einem internen Modell werden Gedächtnisinhalte verstanden, die das Wissen über ein Objekt darstellen. Als perzeptiver Prozess gilt der Vorgang, Reize der Umwelt durch Rezeptoren aufzunehmen, zu kodieren und aktiv zu Ganzheiten zu ordnen. Als kognitive Prozesse werden Vorgänge wie Kennen, Wiedererkennen, Erkennen oder Lernen verstanden, bei denen interne Modelle gespeichert, zugeordnet, angelegt oder verändert werden.

Der Einteilung perzeptiv/kognitiv entsprechend,wird bei der Szenenanalyse im Computer zwischen niederen und höheren Prozessen unterschieden. Auch hier ist die Schnittstelle schwer zu definieren. Hauptaufgabe der niederen Prozesse sind die Segmentation und die Merkmalsextraktion. Die Segmentation ist die Zerlegung der Szene in einfachere Strukturen, die auf verschiedenen Ebenen erfolgen kann. Segmentierte Regionen lassen sich durch Merkmale charakterisieren. Hierbei kann es sich um Merkmale der Form oder der Grauwertverteilung (Texturmerkmale) handeln, mit deren Hilfe höhere Prozesse den segmentierten Regionen interne Modelle zuordnen, bestehende Modelle korrigieren oder neue Modelle konstruieren. Der Vorgang der Zuordnung interner Modelle soll als Analyse bezeichnet werden.

In technischen Systemen können verschiedene Arten von Bildern vorliegen: Binärbilder, Grauwertbilder, Farbbilder, Stereobilder (mit Einschränkungen) und Lichtschnittbilder. Verfahren der Szenenanalyse verarbeiten

vorwiegend Grauwert- und Farbbilder. In dieser Arbeit bleiben Probleme des Farbensehens unberücksichtigt. Im Zusammenhang mit Szenenanalysen im Computer wird hier das auszuwertende Bild stets als diskrete Grauwertmatrix $G(x_i,y_j)$ aufgefaßt, wobei (x_i,y_j) die Koordinaten eines diskreten Bildpunktes sind. Zur Vereinfachung der Notation werden in der gesamten Arbeit die Indizes i und j überall weggelassen, wo sie nicht zum Verständnis notwendig sind. Aufgabe der Szenenanalyse ist die Verarbeitung der Grauwertmatrix, um Aussagen über die in der Szene enthaltenen Objekte und ihre Beziehungen untereinander zu gewinnen und durch interne Modelle zu beschreiben. Fragen der Repräsentation spielen dabei eine wichtige Rolle: bei niederen Prozessen werden das Bild und die Zwischenergebnisse, bei höheren Prozessen die internen Modelle und Szenenbeschreibungen durch geeignete Datenstrukturen repräsentiert.

1.2 ZIELSETZUNG UND BEITRAG DER ARBEIT

Die vorliegende Arbeit hat ihren Ursprung in einer Untersuchung zur "Modellbildung des menschlichen Erkennungssystems" /22/, für die das visuelle System als Gegenstand der Untersuchung ausgewählt wurde. Thematisch ist die Arbeit eingeordnet in das Gebiet der Mustererkennung, in der Probleme der Szenenanalyse im Computer untersucht werden. Phänomene und Leistungen des menschlichen visuellen Systems bilden dabei einen gedanklichen Rahmen, an dem sich die Entwicklung geeigneter Verfahren orientieren kann.

Die Arbeit hat mehrere Ziele. Zuerst wird in einer Literaturdurchsicht Charakteristisches über Sehen zusammengetragen und geordnet. Dabei wird eine knappe Darstellung angestrebt, die Spekulatives vermeidet und sowohl beim menschlichen Sehen als auch bei den Verfahren der Szenenanalyse im Computer die wesentlichen Prinzipien herausarbeitet.

Hauptziel der Arbeit ist die Entwicklung leistungsfähiger niederer Prozesse für Szenenanalysen von natürlichen Szenen. Reine Vorverarbeitungstechniken (wie z.B. Kontrastverbesserung, Rauschunterdrückung,..) werden ebensowenig wie die höheren Prozesse berücksichtigt. Eine Kritik einiger

Verfahren der Literatur, die sich aus Experimenten und theoretischen Überlegungen ergibt, mündet in einen Katalog von Verarbeitungsprinzipien.

Unter Berücksichtigung dieser Prinzipien wird ein neuartiges Konzept entwickelt. Der Grundgedanke dieses Ansatzes besteht darin, die Grauwertmatrix in eine symbolische Datenstruktur zu überführen, noch bevor irgend eine Auswertung vorgenommen wird. Diese Datenstruktur erleichtert Auswertungen, indem sie Kontextinformation in einer hierarchischen Form bereitstellt. Technisch wird der Ansatz durch eine Baumstruktur verwirklicht, durch die das Bild zeilenweise in eine strukturelle Beschreibung von Intensitätsprofilen übergeführt wird. Die anschließende Verarbeitung umfaßt neue Verfahren der Kantendetektion, Flächenbildung, Texturanalyse und der Verfolgung von Teilstrukturen quer zu den Zeilen (Profilverfolgung). Diese Verfahren greifen nur auf die Baumstrukturen zu und nicht mehr auf das Bild. Für diese Auswertung werden Algorithmen entwickelt und experimentell abgestützt. Die entwickelten Techniken werden zusammen mit anderen Verfahren in ein theoretisches Modell niederer Prozesse für Szenenanalysen integriert.

Somit enthält die Arbeit folgende Beiträge:

- Eine vergleichende Erprobung und Kritik einiger bekannter Verfahren zusammen mit der Entwicklung eines eigenen Kantendetektors;
- einen Katalog von Verarbeitungsprinzipien;
- ein neuartiges Konzept für die Repräsentation von Grauwertbildern;
- die Entwicklung von Auswertungstechniken für Kantendetektion, Flächenbildung, Texturanalysen und Profilverfolgung;
- Experimente zur Erprobung dieser Techniken;
- ein theoretisches Modell niederer Prozesse für Szenenanalysen natürlicher Szenen.

2. SZENENANALYSE BEIM MENSCHEN

Die Beiträge des menschlichen Sehsinnes für Szenenanalysen werden hier umrissen, ohne auf Farben- oder Bewegungssehen einzugehen. Man hat zwei Hirnregionen gefunden, die für die Verarbeitung visueller Reize verantwortlich sind: den visuellen Cortex und den Colliculus Superior im Mittelhirn /15/. Während der entwicklungsgeschichtlich jüngere Cortex ausschließlich visuelle Reize verarbeitet, werden im Colliculus auch Reize erfaßt, die von anderen Hirnregionen stammen.

Abbildung 1 zeigt ein Blockdiagramm des visuellen Systems. Die Rezeptoren der Retina wandeln Lichtreize in neuronale Impulse um. Die Auflösung des Rezeptorfeldes ist hoch in der Zone des schärfsten Sehens, der Fovea, und fällt zur Peripherie hin ab.Von den Rezeptorzellen wird die Lichtinformation über den optischen Nerv zum Corpus Geniculatum Laterale (CGL) und zum Colliculus Superior übertragen. Durch Querverschaltungen werden die Ein-Ausgabe-Ebenen von Zellen mittels Erregung und Hem-

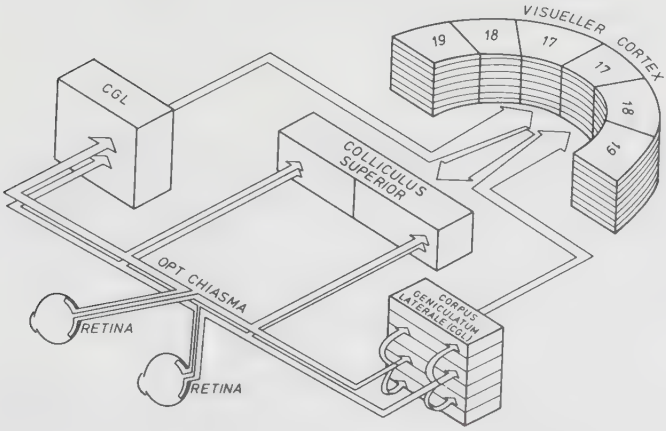


ABBILDUNG 1 : BLOCKDIAGRAMM DES MENSCHLICHEN VISUELLEN SYSTEMS

mung durch Nachbarzellen beeinflußt. Alle Zellen - bis hin zur Retina - reagieren nur, wenn ein Reiz innerhalb bestimmter Bereiche der Retina auftritt. Derjenige Bereich der Retina, innerhalb dessen ein Lichtreiz eine Reaktion einer Zelle auslöst, wird "rezeptives Feld" der betreffenden Zelle genannt /70/.

Im CGL werden die beiden Retinabilder in sechs übereinanderliegende Schichten zusammengeführt, die abwechselnd von jeweils einem Auge betrieben werden. Die Verschmelzung der beiden Retinabilder findet jedoch erst im visuellen Cortex statt. Dieser gliedert sich in die drei Bereiche der Area 17, 18 und 19. In allen drei Bereichen bleibt die Topologie der Retina erhalten. Neben radialsymmetrischen rezeptiven Feldern werden hier Felder gefunden, die bevorzugt auf Linien bestimmter Neigung oder auf geometrische Merkmale reagieren - z.B. auf Linienenden, Knicke oder Linienunterbrechungen. Man unterscheidet einfache, komplexe und hyperkomplexe rezeptive Felder, die als Detektoren für einfache Formmerkmale wirken /64/. Wie diese Formmerkmale weiterverarbeitet werden, ist nicht bekannt. Es ist unwahrscheinlich, daß auch höhere Formen durch feste Detektoren extrahiert werden. Vielmehr wird angenommen, daß die einfachen Formmerkmale durch eine Grammatik zu höheren Formen verknüpft werden /38/.

Im Colliculus Superior wird eine wichtige Schaltstelle zur Steuerung der Augenbewegungen vermutet. Auch hier bleibt die Topologie der Retina erhalten, allerdings wird die Fovea nicht durch eine hohe Auflösung bevorzugt. Es gibt Hinweise dafür, daß der Colliculus eine Hirnregion ist, die zur Lokalisation von Objekten dient - und zwar unabhängig von einer Formanalyse / 1 /. Es wird vermutet, daß im Cortex schwerpunktmäßig die Analyse von Formen vollzogen wird, während im Colliculus die Lage, Größe und zeitliche Veränderung von Objekten ermittelt werden.

Neben den Reizmustern der Retina liegen im Hirn noch okulomotorische Reize vor. Jedes Auge wird durch sechs Muskeln bewegt und führt eine Vielzahl von Bewegungen durch, von denen für die Szenenanalyse zwei Arten von Bedeutung scheinen: Makro-Sakkaden und Vergenzbewegungen. Makro-Sakkaden sind schnelle Bewegungen, die die Fovea über dem Seh-

feld verschieben. Untersuchungen von Augenbewegungen beim Betrachten von Bildern zeigen, daß eine Szene nicht vollständig abgefahren wird. Es werden lediglich Gebiete betrachtet, die relevante Informationen enthalten. Da Relevanz der Information von der Aufgabenstellung und den Intentionen des Beobachters abhängt, muß dieser zumindest Hypothesen zur Verfügung haben, wo interessante und relevante Information im Sehfeld gefunden werden kann. Augenbewegungen werden vermutlich durch interne Modelle und daraus resultierende Erwartungen und Hypothesen ebenso gesteuert wie durch Auffälligkeiten in einer Szene /51, 54, 73, 76, 80/.

Vergenzbewegungen verändern den Winkel, der durch die Sehachsen der beiden Augen gebildet wird, und werden durch die Disparität der Retinabilder ausgelöst. Disparität ist ein Maß für die gegenseitige Verschiebung von Punkten auf den Retinae, die zum gleichen Objektpunkt gehören. Vergenzwinkel und Disparität sind die binokularen Merkmale zur Berechnung der Raumtiefe und werden durch folgende monokulare Tiefenmerkmale ergänzt: Überdeckungen, wahrgenommene Größe bekannter Objekte, Bewegungsparallaxe, lineare Perspektive, Dunstperspektive, Texturgradient, Licht und Schatten sowie Akkomodation /55/. Aus all diesen Merkmalen wird die Tiefe von Objekten im Raum bestimmt und dadurch eine Segmentation in räumlich getrennte Objekte vorgenommen /26/. Es gibt Befunde, die zeigen, daß es neuronale Mechanismen zur Auswertung der Disparität gibt / 2 /.

Neben der Tiefenwahrnehmung gibt es zwei weitere Mechanismen der Segmentation, die auf Farbe und Textur aufbauen. Da sich Texturen durch Kombination vieler Reize erzeugen lassen, ist es nicht plausibel, anzunehmen, daß es auch für Texturen spezifische Detektoren gibt. Die Fähigkeit, Texturen wahrzunehmen und zu unterscheiden, scheint im wesentlichen auf der örtlichen Verteilung einfachster Merkmale aufzubauen, wobei diese Merkmale nach Durchmesser, Weite, Länge und Orientierung gruppiert werden. Modelle der Texturwahrnehmung bauen deshalb auf konstruktiven Mechanismen auf, bei denen die Relationen von Bildelementen zueinander dazu führen, daß Bildelemente zu perzeptiven Gruppen zusammengefaßt werden. Die Gestaltstheorie hat Gruppierungsmechanismen auf Gestaltsprinzipien zurückgeführt: Nähe, Ähnlichkeit, Geschlossenheit, Symmetrie, gute

Fortsetzung /38, 80/. Es ist jedoch nicht klar, welche Bedeutung diese Mechanismen für die Analyse natürlicher Szenen haben.

Durch die Segmentationsmechanismen des Tiefen-, Farben- und Textursehens wird eine Trennung in Figur und Hintergrund erreicht. Die Fähigkeit der Trennung von Hintergrund und Figur scheint angeboren und unabhängig vom Formensehen zu sein /27/. Ein besonderes Phänomen ist die Fähigkeit des Menschen, diese Trennung auf verschiedenen Ebenen zu vollziehen: man kann beim Betrachten eines Waldes als Figur auffassen: den Wald, eine Baumgruppe, einen Baum, einen Stamm, Borke, ..., usw. In allen Fällen liegt dasselbe visuelle Muster vor. Diese verschiedenen Betrachtungsweisen stehen dabei nicht nebeneinander, sondern schließen sich gegenseitig aus. Die Existenz dieser verschiedenen Betrachtungsebenen erfordert Mechanismen, die eine Szene so segmentieren, daß das "Richtige" ausgesondert wird.

Über die Speicherung interner Modelle und die Zuordnung von Modellen zu den extrahierten Figuren liegen kaum schlüssige Befunde vor. Es wird u.a. eine Reihe von Gedächtnistypen postuliert, die als hintereinandergeschaltet gedacht werden, wobei die ersten Gedächtnisse sehr viel Information für sehr kurze Zeit, die nachfolgenden immer stärker abstrahierte Information immer länger speichern /52/. Für die Zuordnung werden assoziative Prozesse vermutet, bei denen es genügt, einen Teil eines Musters vorzulegen, um das Gesamtmuster abzurufen.

Zusammenfassend ergibt sich: der Sehvorgang ist von der Verarbeitung der visuellen Reize her ein paralleler Prozeß; durch die Verschiebung der Fovea über dem Sehfeld findet auch ein sequentieller Vorgang statt. Die Verarbeitung findet in verschiedenen Hirnregionen statt, die zum Teil hintereinander, zum Teil nebeneinander geschaltet sind. Ein wesentliches Prinzip ist die Verarbeitung visueller Reize durch Erregung und Hemmung benachbarter Zellen untereinander. Kennzeichnend ist, daß die Topologie der Retina stets erhalten bleibt und daß die Trennung von Hintergrund und Figur unabhängig vom Formensehen und auf verschiedenen Auflösungsebenen erfolgen kann.

3. SZENENANALYSE IM COMPUTER

Im Computer liegt eine Szene als digitale Grauwertmatrix G (x,y) vor. Es ist Aufgabe der Szenenanalyse, das Bild in Regionen zu zerlegen, den Regionen interne Modelle zuzuordnen und so eine Beschreibung der in der Szene enthaltenen Objekte und ihrer Lage zu liefern. Zur Lösung dieser Aufgabe müssen Probleme der Vorverarbeitung, Segmentation, Merkmalsextraktion, Repräsentation und der Analyse gelöst werden. Der Zielsetzung der Arbeit entsprechend (vgl. Abschnitt 1.2), werden Probleme der Vorverarbeitung hier nicht berücksichtigt. Den Definitionen in Abschnitt 1.1 folgend, werden die Verarbeitungsschritte bis zur Segmentation und Merkmalsextraktion als "niedere Prozesse" aufgefaßt, während die Analyse den "höheren" Prozessen zugerechnet wird. Probleme der Repräsentation treten sowohl bei niederen als auch höheren Prozessen auf.

Es ist umstritten, wo die Schnittstelle zwischen niederen und höheren Prozessen liegen sollte. Die ersten Verfahren der Szenenanalyse waren durch eine rein lokale Informationsverarbeitung gekennzeichnet, bei der globale Information und interne Modelle keinen Einfluss auf niedere Prozesse hatten. Die folgende Generation von Verfahren war durch modellgestützte Analysen charakterisiert, bei denen höhere Prozesse frühzeitig in den Segmentationsprozess eingreifen und die niederen Prozesse lenken. Dies setzt jedoch A-priori-Kenntnisse voraus, von denen nicht erwartet werden kann, daß sie in einer allgemeinen und flexiblen Weise für Szenenanalysen verfügbar sind. Deshalb zeichnet sich heute ein Trend ab, autonome niedere Prozesse zu entwickeln, die Daten-gestützt sind, jedoch Entscheidungen nicht mehr allein aufgrund lokaler Information treffen.

3.1 VERFAHREN DER LITERATUR

Es gibt zwei Vorgehensweisen, eine Szene zu segmentieren:

(1) die Ermittlung von Konturlinien;
(2) die Ermittlung von Regionen aufgrund ihrer Textur.

Diese Vorgehensweisen erfassen unterschiedliche Merkmale. Im ersten Fall sind es Merkmale, die Unterschiede zwischen Regionen charakterisieren und damit Konturen festlegen. Im zweiten Fall sind es Merkmale,

die Ähnlichkeiten von Eigenschaften von Bildpunkten bzw. Regionen erfassen. Bei beiden Vorgehensweisen stellen Texturen das schwierigste Problem dar. Die Modellierung, Beschreibung und Analyse von Texturen ist deshalb ein zentrales Problem niederer Prozesse. Allgemein wird Textur aufgefaßt als - geordnetes oder zufällig verteiltes - Auftreten von Texturgrundelementen ('Textural Primitives').

KONTUREN

Verfahren, die mit Konturen arbeiten, gliedern sich in folgende Verarbeitungsschritte: Detektion von Kanten- und Linienelementen; Verfolgung solcher Elemente und Verknüpfung zu Konturlinien (Linienverfolgung); Markierung der Linien durch Marken, die die Bedeutung der Linien in der Szene wiedergeben; Segmentation in Regionen und Analyse dieser Regionen durch Zuordnung interner Modelle.

Für die Kanten- und Liniendetektion liegt eine Vielzahl von Operatoren vor, die z.T. als Hochpaßfilterung im Orts-Frequenzbereich, z.T. als Differentiation im Ortsbereich definiert sind, wobei Verfahren im Ortsbereich überwiegen /13/. Systematische Wege zum Entwurf geeigneter Operatoren existieren ebensowenig wie Gütekriterien zur Beurteilung ihrer Leistungsfähigkeit. Für den Entwurf ist man weitgehend auf Intuition angewiesen; die Beurteilung erfolgt durch Inspektion verarbeiteter Bilder. Allen Operatoren liegt ein Modell dessen zugrunde, was als Kante oder Linie aufgefaßt wird. Abbildung 2 zeigt drei Intensitätsverläufe, die als typische Querschnitte an Kanten und Linien in Szenen gefunden wurden, die aus einfachen Polyedern bestehen /36/.

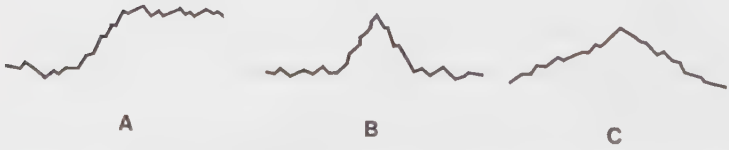


ABBILDUNG 2 : INTENSITÄTSPROFILE VON KANTEN UND LINIEN IN BILDERN AUS DER "BLOCKSWELT" ('BLOCK'S WORLD')

A - STUFE B - SPITZE C - DACH (NACH /36/)

Implizite Modelle liegen den meisten der linearen und nicht-linearen Detektoren zugrunde, die auf einem Fenster definiert sind, welches über dem Bild verschoben wird /9 , 16 , 59 , 60 /. Hier ist das Modell in der Gewichtungsfunktion und der Rechenvorschrift der Operatoren enthalten. Werden explizite Modelle formal dargestellt, so lassen sich Operatoren verwirklichen, die bezüglich vorgegebener Maße optimal sind /37/.

Die Aufgabe der Linienverfolgung besteht darin, aus der Vielzahl der Punkte, die die Kantendetektion liefert, diejenigen auszuwählen, die tatsächlich auf einer Kante bzw. Linie liegen, und sie zu Konturlinien zu verknüpfen. Hierfür werden eingesetzt: (1) lokale oder (2) globale Verfahren. Das einfachste lokale Verfahren ist eine Schwellwertoperation, bei der alle Punkte unterdrückt werden, deren Operatorwert unter einem vorgegebenen Schwellwert liegt. Ein geeigneter Schwellwert wird aus Histogrammanalysen abgelesen /11/. Derartige lokale Methoden bewähren sich selbst in einfachen Szenen nicht. Eine Verbesserung wird durch Berücksichtigung globaler Information erzielt. Hierfür liegen u.a. folgende Techniken vor:

(1) Iterative Verfahren, bei denen Zwischenergebnisse berücksichtigt werden. Beispiele sind die heuristische Suche /46/ oder dynamisches Programmieren /20, 42/.
(2) Das Ausnützen von A-priori-Information, wobei interne Modelle die Linienverfolgung lenken /65/.
(3) Das Erstellen von Plänen, die aus Bildrepräsentationen niedriger Auflösung gewonnen werden und die Grobstrukturen eines Bildes enthalten. Die Grobstrukturen liefern globale Information, mit deren Hilfe die Linienverfolgung gelenkt wird / 29 , 39 , 66, 68 /.

Die gefundenen Linien müssen anschließend durch Kurven beschrieben werden, z.B. indem man Ausgleichsgeraden durch die Linienpunkte legt /50/.

In den derart gewonnenen Liniendarstellungen müssen diejenigen Regionen gefunden werden, die den Objekten in der Szene entsprechen. Dies geschieht durch die Markierung von Ecken /28/ und Linien /12, 69/ mit Marken, die Aufschluß über Eckengeometrie, Orientierung von Flächen, Art der Beleuchtung und Beleuchtungsrichtung geben. Der wesentliche Grundgedanke dieser Methode besteht darin, nach einer Initialzuweisung die Marken solange iterativ zu ändern, bis eine konsistente Markierung

vorliegt. Der iterative Austausch von Marken ist durch Relaxationsprozeduren formalisiert worden /61/. Für Probleme der Linienverfolgung und Markierung wurden Relaxationsprozeduren bereits erfolgreich angewandt /57, 63, 78, 79/.

REGIONEN

Verfahren, die nur Konturlinien erfassen, liefern keine Information über Eigenschaften der Grauwertverteilung innerhalb der ermittelten Regionen. Da in natürlichen Szenen Objekte oft besser durch Oberflächeneigenschaften als durch ihre Form definiert sind, bleibt eine wichtige Information quelle - nämlich Texturen - unberücksichtigt. In dem Ausmaß, mit dem sich das Interesse von der Analyse einfacher Laborszenen auf die Analyse natürlicher Szenen verlagerte, setzte auch die Entwicklung von Verfahren zur Flächenbildung ('Region Growing') ein.

Verfahren der Flächenbildung nehmen eine Segmentation des Bildes in Zusammenhangskomponenten vor, auf die vorgegebene Prädikate zutreffen. Die Verfahren lassen sich durch folgende Kriterien kennzeichnen:

(1) Die Verarbeitungsrichtung:
   (a) einzelne Bildpunkte oder kleine Regionen werden zu immer größeren Regionen verschmolzen;
   (b) das Bild oder große Regionen werden immer weiter aufgespalten.

(2) Das Verarbeitungskriterium:
   (a) lokale Merkmale innerhalb von Regionen oder zwischen Regionen werden ausgewertet;
   (b) globale Merkmale - z.B. die Verteilung der Grauwerte im gesamten Bild - werden ausgewertet.

(3) Die Art der Daten:
   (a) die ursprünglichen Grauwerte;
   (b) abgeleitetet Daten wie Filterwerte, Frequenzen, Koeffizienten numerischer Approximation, usw.;
   (c) semantische Daten wie symbolische Marken.

(4) Die Initialsegmentation:
   (a) zufällig oder willkürlich in beliebige Regionen;
   (b) deterministisch in ein Raster;
   (c) merkmalsabhängig - z.B. in Regionen gleichen Grauwerts.

Die ersten Verfahren zur Flächenbildung /6, 8, 49/ sind gekennzeichnet durch das Verschmelzen kleiner Regionen aufgrund lokaler Merkmale, wobei mit den ursprünglichen Grauwerten gearbeitet wird. Die Initialsegmentation erfolgt deterministisch oder merkmalsabhängig.

Eine stärkere Ausnutzung globaler Information wird durch die Einführung hierarchischer Repräsentationen und die Ergänzung des "Schmelzens" durch "Spalten" erreicht. Als hierarchische Repräsentation wird häufig ein Pyramidensystem eingesetzt, in dem ein Bild auf mehreren übereinanderliegenden Ebenen gespeichert wird /29, 66/. "Spalten und Schmelzen" ('Split & Merge') sind durch Übergänge zwischen verschiedenen Auflösungsebenen definiert. Als Daten werden Koeffizienten einer numerischen Approximation des Bildes /56/ oder Farbinformationen und Grauwerte /31/ verarbeitet. Auch Histogramme liefern globale Information. Es werden eindimensionale Histogrammme von Merkmalen untersucht /53, 67/ oder zweidimensionale Histogramme von Merkmalspaaren als Pseudobilder in einem Pyramidensystem ausgewertet /31/.

Bei den bisher beschriebenen Verfahren werden Grauwerte oder abgeleitete Daten zur Flächenbildung ausgewertet. Der Einsatz semantischer Marken, die vorläufige Analysen segmentierter Regionen darstellen, erlaubt es, Techniken zu entwickeln, bei denen die Segmentation zusammen mit der Suche nach einer konsistenten Markierung abläuft /74, 75/.

Neben der Entwicklung geeigneter Segmentationsverfahren stellt die Erfassung und Beschreibung von Texturen ein weiteres wichtiges Problem niederer Prozesse dar. Beschreibungen werden im Ortsbereich und im Frequenzbereich vorgenommen. Im Ortsbereich arbeitet eine Reihe von Verfahren, die die Häufigkeit von Grauwertübergängen auswerten. Es werden die Häufigkeiten ermittelt von:

(1) Kanten- und Linienelementen oder Flecken in einer vorgegebenen Region /53, 77/;
(2) Strecken mit konstantem Grauwert ("Lauflängen") /25/;
(3) Grauwertpaaren (i,j) /14, 33/;
(4) lokalen Extremwerten vorgegebener Intensität in rechteckigen Bildausschnitten /10, 48/.

Einer der wenigen Ansätze der Texturbeschreibung im Frequenzbereich wertet Fourier-Leistungsdichte-Spektren aus /5/, aus denen sich ins-

besonderedie Orientierung von Texturgrundelementen ablesen läßt.

REPRÄSENTATIONEN

Im Gegensatz zu Verfahren, die Texturbeschreibungen direkt aus dem Grauwertbild gewinnen, steht ein Ansatz, dessen Grundgedanke darauf beruht, das Bild zuerst in eine symbolische Repräsentation zu überführen und erst aus dieser Repräsentation Texturen zu extrahieren /43, 44/. Diese Repräsentation wird "Erst-Skizze" ('Primal Sketch') genannt und durch Faltung des Bildes mit Operatoren verschiedener Größe und Gewichtungsfunktion gewonnen. Aus den Ergebnissen der Faltung wird ein symbolisches Vokabular gewonnen, das Grauwertänderungen beschreibt und Information über Richtung, Größe, Intensität, Lage und Endpunkte der Elemente der Erst-Skizze enthält. Aus diesen Informationen wird Texturinformation gewonnen. In einer fortführenden Arbeit /45/ wird die Bedeutung von Repräsentationen für die Szenenanalyse herausgestellt: Szenenanalyse kann nicht durch eine einzige, drastische Transformation des Bildes in eine Szenenbeschreibung erfolgen, sondern muß über eine Reihe von Zwischenergebnissen ablaufen, die in jeweils problemspezifischen Repräsentationen dargestellt werden. An die "Erst-Skizze" schließt deshalb eine Beobachter-orientierte "2 1/2-D-Skizze" (' 2 1/2-D-Sketch') an, aus der eine objektbezogene Beschreibung - die "3-D-Modell-Repräsentation" ('3-D-model-representation') gewonnen wird.

Ein äußerst komplexes System für Szenenanalysen, das VISIONS-System /30, 32/, weist ebenfalls mehrere Repräsentationen auf: eine Szene wird in einer Pyramidenstruktur ('Cone' genannt) gespeichert und durch niedere Prozesse verarbeitet. Die Ergebnisse werden in einem Graphen gespeichert, der in drei Schichten Regionen, Linien-Segmente und Endpunkte enthält und deshalb RSE-Struktur genannt wird. Die RSE-Struktur stellt die Schnittstelle zwischen niederen und höheren Prozessen dar. Repräsentationen für höhere Prozesse befinden sich im sogenannten Modell-Bildner, in dem sowohl interne Modelle als auch die aktuelle Szenenbeschreibung enthalten sind.

Probleme der Repräsentation treten sowohl bei niederen als auch bei höheren Prozessen auf. Für Repräsentationen, die von niederen Prozessen verarbeitet werden, wurden im wesentlichen zwei Arten von Datenstrukturen entwickelt: (1) die bereits erwähnten Pyramidensysteme und (2)

verschiedene Baumstrukturen. Ein Verfahren, eine Baumstruktur zu gewinnen, besteht in einer wiederholten Zerlegung von Quadraten eines Bildes bis Quadrate mit homogener Grauwertverteilung erreicht sind /40/. Der Ablauf dieser Zerlegung wird in einer Baumstruktur festgehalten. Eine Zerlegung des Bildes, die ebenfalls in einer Baumstruktur festgehalten wird, erfolgt in /41/ durch Schwellwertoperationen für alle diskreten Grauwerte und Feststellung der Konturlinien von Zusammenhangskomponenten. Flexibler als diese beiden Ansätze ist das Konzept des "Affinität-Baumes" ('Affinity Tree'), bei dem benachbarte Regionen, deren Grauwertverteilungen eine gewisse Affinität zueinander haben, zu einem Knoten im Baum zusammengefaßt werden /24/.

Repräsentationen, die im Zusammenhang mit höheren Prozessen stehen, wurden oben bereits erwähnt. Die Repräsentation interner Modelle wird in großen Einheiten vorgenommen, die als Stereotype von Situationen, Objekten, usw. auftreten. Diese Einheiten sind stark strukturiert und stellen eine Vielzahl von Information zur Verfügung, die insbesondere die Zuordnung eines Modelles zu Regionen im Bild unterstützen /3, 4, 47/.

3.2 EXPERIMENTELLE ERPROBUNG UND KRITIK AUSGEWÄHLTER VERFAHREN

In diesem Abschnitt werden zunächst Experimente beschrieben, in denen einige Verfahren der Literatur für Kompression, Kantendetektion und Flächenbildung sowie ein eigener Kantendetektor untersucht wurden. Ziel der Experimente war die Untersuchung der Leistungsfähigkeit niederer Prozesse für die Segmentation und Merkmalsextraktion. Im Verlauf dieser Experimente zeigten sich grundsätzliche Schwächen der Verfahren. Deshalb mündet dieser Abschnitt in eine Kritik, die auf den Ergebnissen der Experimente und auf theoretischen Einsichten aufbaut.

Fast alle Experimente wurden mit Bildern durchgeführt, die 256 x 256 Bildpunkte und 256 Grauwerte haben. Grauwert o entspricht Schwarz, Grauwert 255 entspricht Weiß. Ursprung der Bildpunkt-Koordinaten ist die linke obere Ecke der Grauwertmatrix. Eine X-Koordinate x_{i+1} liegt rechts von x_i , eine Y-Koordinate y_{j+1} liegt unterhalb von y_j.

KOMPRESSION

Die im Abschnitt 3.1 beschriebenen Pyramidensysteme haben als hierachische Repräsentation von Bildern Verbreitung gefunden. Ihre Konstruktion erfolgt durch mehrfache Kompression im Ortsbereich (Abbildung 3).

Üblicherweise werden die Auflösungsebenen durch Anwendung eines Operators erzeugt, der auf einem Fenster definiert ist, welches N x M zusammenhängende Bildpunkte bedeckt und nicht-überlappend auf dem Bild verschoben wird. Innerhalb des Fensters wird der Durchschnitt D

$$D\,(x_i,\, y_j) = \frac{1}{N \cdot M} \sum_{i=1}^{N} \sum_{j=1}^{M} (x_i, y_j) \qquad (3.1)$$

der Grauwerte der vom Fenster bedeckten Bildpunkte berechnet und als neuer Grauwert demjenigen Bildpunkt zugewiesen, der auf der (k+1)-ten

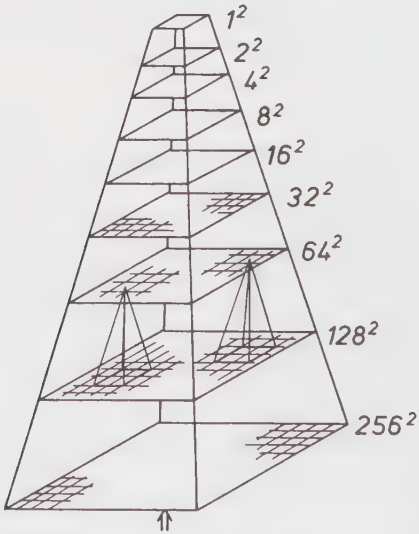


ABBILDUNG 3 : BEISPIEL FÜR EIN PYRAMIDENSYSTEM ZUR HIERARCHISCHEN REPRÄSENTATION EINES BILDES

DAS BILD LIEGT AUF DER UNTERSTEN EBENE MIT DER URSPRÜNGLICHEN AUFLÖSUNG (256 X 256 BILDPUNKTE) VOR. DURCH NICHT-ÜBERLAPPENDE ANWENDUNG EINES OPERATORS WERDEN JEWEILS 2 X 2 PUNKTE EINER EBENE ZU EINEM PUNKT DER NÄCHSTHÖHEREN EBENE ZUSAMMENGEFASST.

Ebene dem Fenster auf der k-ten Ebene entspricht. Als Alternativen können andere Rechenvorschriften - z.B. Minimum, Maximum oder Modalwert - als Wert des Operators eingesetzt werden.

Pyramidensysteme werden insbesondere verwendet, um "Pläne" zu erzeugen, mit deren Hilfe niedere Prozesse gelenkt werden. "Pläne" enthalten Grobstrukturen, die auf Ebenen geringer Auflösung ermittelt wurden. Diese Strukturen werden "nach unten" in Ebenen hoher Auflösung projiziert. Um die Güte solcher "Pläne" zu untersuchen, wird die Wirkung der Projektion von Grobstrukturen auf das Bild gezeigt, indem diese Grobstrukturen mit der ursprünglichen Auflösung des Bildes dargestellt werden. Abbildung 4 demonstriert, wie schwierig es ist, "Pläne" dieser Art für eine sorgfältige Analyse zu verwenden.

ABBILDUNG 4 : WIRKUNG DER KOMPRESSION MIT OPERATOR D (GL. (3.1))
A - ORIGINAL (900 X 1000 BILDPUNKTE , 256 GRAUWERTE)
B - H : PROJEKTION DER KOMPRIMIERTEN BILDER AUF ORIGINALGRÖSSE
FENSTERDIMENSIONEN (N X M) : B - 2x2 C - 4x4 D - 8x8
E - 16x16 F - 32x32 G - 50x50 H - 100x100

Aus Abbildung 4 H kann bestenfalls auf eine dunkle Region in der Bildmitte geschlossen werden. Aus den Block-Strukturen kann weder über die Form noch die exakte Plazierung der Region eine Hypothese formuliert werden, die brauchbare Information für eine sorgfältige Analyse liefert.

Die Wirkung dieser komprimierenden Operatoren hängt entscheidend davon ab, wie groß die bedeutungsvollen Regionen eines Bildes im Verhältnis zur Fenstergröße sind und wie die Fenster bezüglich den Grenzen zwischen benachbarten Regionen plaziert sind. Die starre Vorgabe der Pyramidenstruktur und des Auflösungsrasters erweist sich als Ursache grundsätzlicher Schwierigkeiten, die diesen Ansatz als nicht befriedigend ausweisen.

KANTENDETEKTION

Für die Erprobung von Segmentationsverfahren wurden zwei Techniken ausgewählt:
(1) Kantendetektion durch lokale Operatoren und (2) Segmentation durch Histogrammanalysen. Für Verfahren der Kantendetektion wurden zwei Operatoren ausgewählt, die exemplarisch für eine Klasse von Verfahren sind; das "Robert's Kreuz" ('Robert's Cross' /16/) und der "Sobel-Operator" /16/. Beide Operatoren sind auf quadratischen Fenstern N x N definiert (Abbildung 5 A, B). Das "Robert's Kreuz" ist auf einem 2 x 2 -Fenster definiert, für den "Sobel-Operator" ist N=3. Neben diesen beiden Operatoren wurde ein eigener Operator erprobt /23/. Das "Robert's Kreuz" R (x,y) wird häufig in einer rechnerisch einfachen Version implementiert:

$$R(x_i,y_j)=|G(x_i,y_j)-G(x_{i+1},y_{j+1})|+|G(x_{i+1},y_j)-G(x_i,y_{j+1})| \quad (3.2)$$

Dabei kennzeichnet (x_i,y_j) die Position des Fensters. Der "Sobel-Operator" $S(x_i,y_j)$ ist definiert durch:

$$S(x_i,y_j) = [S_x^2 + S_y^2]^{1/2} \quad (3.3)$$

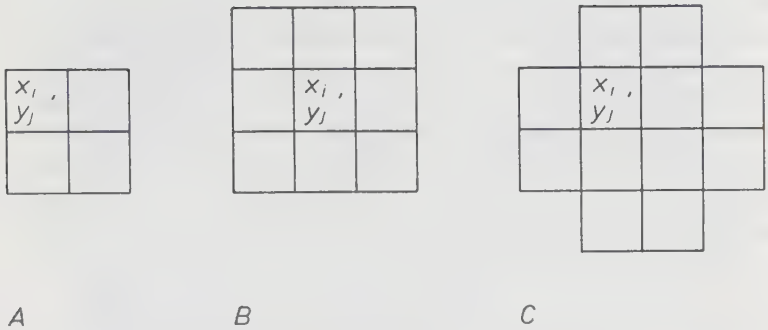


ABBILDUNG 5 : FENSTER LOKALER OPERATOREN ZUR KANTENDETEKTION
A - $R(X_I, Y_J)$ B - $S(X_I, Y_J)$ C - $F(X_I, Y_J)$

wobei:

$$S_x = [G(x_{i+1}, y_{j-1}) + 2 \cdot G(x_{i+1}, y_j) + G(x_{i+1}, y_{j+1})] - [G(x_{i-1}, y_{j-1}) + 2 \cdot G(x_{i-1}, y_j) + G(x_{i-1}, y_{j+1})] \quad (3.4)$$

und:

$$S_y = [G(x_{i-1}, y_{j-1}) + 2 \cdot G(x_i, y_{j-1}) + G(x_{i+1}, y_{j-1})] - [G(x_{i-1}, y_{j+1}) + 2 \cdot G(x_i, y_{j+1}) + G(x_{i+1}, y_{j+1})] \quad (3.5)$$

Während der "Sobel-Operator" die partiellen Ableitungen in X- und Y-Richtung approximiert, berücksichtigt der eigene Operator $F(x_i, y_j)$ durch die Form seines Fensters (siehe Abbildung 5 C) auch die gemischten Glieder, die den beiden Diagonalen entsprechen.

Der Operatorwert ergibt sich aus:

$$F(x_i, y_j) = \max(LDG, RDG, HOR, VRT) \quad (3.6)$$

$$LDG = |[G(x_{i-1}, y_j) + G(x_i, y_{j-1})]| - |[G(x_{i+2}, y_{j+1}) + G(x_{i+1}, y_{j+2})]| \quad (3.7)$$

$$RDG = |[G(x_{i+1}, y_{j-1}) + G(x_{i+2}, y_j)]| - |[((x_{i-1}, y_{j+1}) + G(x_i, y_{j+2})]| \quad (3.8)$$

$$HOR = |[G(x_{i-1}, y_j) + G(x_{i-1}, y_{j+1})]| - |[G(x_{i+2}, y_j) + G(x_{i+2}, y_{j+1})]| \quad (3.9)$$

$$VRT = |[G(x_i, y_{j-1}) + G(x_{i+1}, y_{j-1})]| - |[G(x_i, y_{j+2}) + G(x_{i+1}, y_{j+2})]| \quad (3.10)$$

Durch die Bestimmung des Maximums erhält man gleichzeitig Information über die ungefähre Richtung des Kantenelementes in einer von 4 Neigungen. Aus $S(x_i,y_j)$ können dagegen nur 2 Richtungen ermittelt werden, indem analog das Maximum aus S_x, S_y berechnet wird. Durch eine Abfrage nach dem Vorzeichen kann außerdem festgestellt werden, auf welcher Seite des Fensters die Kante plaziert ist. Damit liefert dieser Operator mehr Information als die beiden Operatoren R und S, benötigt allerdings auch mehr Rechenaufwand als diese. (Das Verhältnis des Rechenaufwandes der Operatoren R : S : F liegt ungefähr bei 1 : 2 : 3).

Nach Anwendung dieser Operatoren liegen "Gradientenbilder" vor, deren Grauwerte den Operatorwert wiedergeben. Die gewünschte Kanteninformation erhält man durch Auswahl der Punkte, die in Gebieten hoher Grauwertänderungen liegen. Dies wird durch eine Schwellwertoperation erreicht, durch die Punkte geringer Grauwertänderung unterdrückt werden.

Abbildung 6 zeigt die Wirkung der Operatoren R, S und F. Durch R werden nicht nur echte Kantenpunkte detektiert: eine Vielzahl von Bildpunkten, die nicht auf Kanten liegen, weist Operatorwerte größer Null auf (vgl. Abb. 6 B).Nach Unterdrückung der Störpunkte liegen jedoch unvollständige Konturen vor (vgl. Abb. 6 C). Die Wirkung der Operatoren S und F ist besser, da ihre Fenster größer sind, und die Operatoren deshalb kleinen Grauwertänderungen gegenüber unempfindlich sind. Beim Betrachten dieser Gradientenbilder läßt das menschliche Sehsystem den Betrachter Linien sehen, obwohl im Computer erst eine Matrix mit Kantenpunkten vorliegt, die durch die Linienverfolgung miteinander verbunden werden müssen. Die Vielzahl der zu verknüpfenden Punkte macht dies zu einer äußerst schwierigen Aufgabe, da diese Kantendetektoren (bis auf die Richtungsinformation) keine Information liefern, die die Linienverfolgung unterstützt. Die Experimente machen deutlich, wie sehr die Ergebnisse der Kantendetektion von der Form, Größe und Lage der Fenster abhängen, auf denen die Operatoren definiert sind.

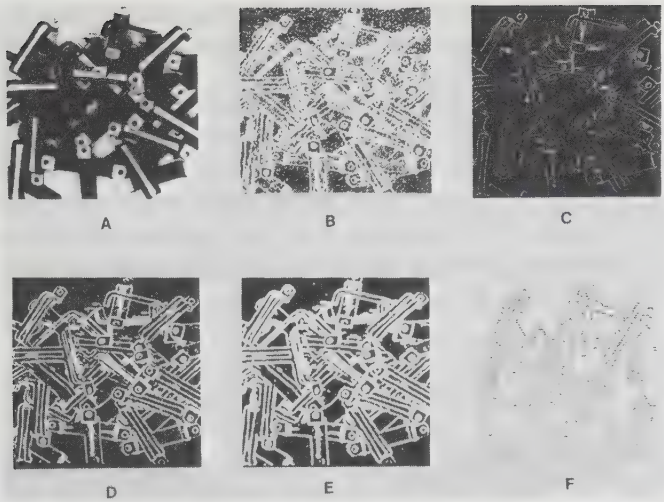


ABBILDUNG 6 : KANTENDETEKTION DURCH LOKALE OPERATOREN

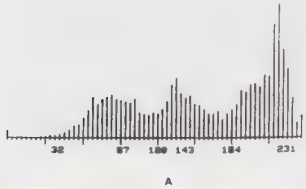
A - GRAUWERTBILD (256 GRAUWERTE)
B - GRADIENTENBILD NACH ROBERTS KREUZ
C - GRADIENTENBILD NACH ROBERTS KREUZ UND SCHWELLWERT-OPERATION (SCHWELLWERT T=40)
D - GRADIENTENBILD NACH SOBEL-OPERATOR UND SCHWELLWERTOPERATION (T=40)
E - GRADIENTENBILD NACH F-OPERATOR UND SCHWELLWERTOPERATION (T=40)
F - RICHTUNGSMATRIX DES F-OPERATORS (AUSSCHNITT : OBERE HÄLFTE VON ABB. 6 E)

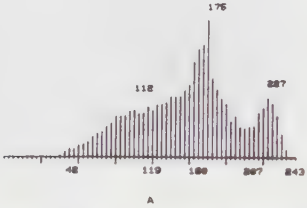
SEGMENTATION DURCH HISTOGRAMMANALYSEN

Zu einem ähnlichen Ergebnis führen auch die Experimente mit Histogrammanalysen zur Flächenbildung. Die Grundidee derartiger Histogrammanalysen beruht auf der Hypothese, daß Merkmale, die Regionen im Bild definieren, in Histogrammen zu Modalwerten führen. Deshalb werden in den Histogrammen (im einfachsten Fall sind dies Histogramme der Grauwerte) Modalwerte gesucht. Aus ihnen werden Wertebereiche für die Merkmale der gesuchten Regionen ermittelt. Die Trennung in Wertebereiche erfolgt an den relativen Minima zwischen den Modalwerten. Durch entsprechende Schwellwerte wird anschließend versucht, im Bild die gesuchten Regionen zu finden. Dabei definiert eine untere und eine obere Schwelle eine "Grauwertscheibe". Nur Bildpunkte, deren Grauwert innerhalb dieser Grau-

wertscheibe liegen, werden berücksichtigt. Diese Punkte werden mit Grauwert 255 (Weiß) dargestellt, alle anderen mit Grauwert o (Schwarz).

Die Abbildungen 7 und 8 zeigen Beispiele des Verfahrens. Das einfachere Bild aus Abbildung 7 weist in seinem Histogramm drei Modalwerte auf (Abb. 7 A), die von den dunkeln Regionen der Kondensatoren, den mittelgrauen Regionen der Kondensatoren und den hellen Regionen des Hintergrundes stammen. Die Verarbeitung des Bildes in Grauwertscheiben mit den entsprechenden Wertebereichen erfaßt damit in jeder Grauwertscheibe die richtigen Regionen (vgl. Abb. 7 C - E).





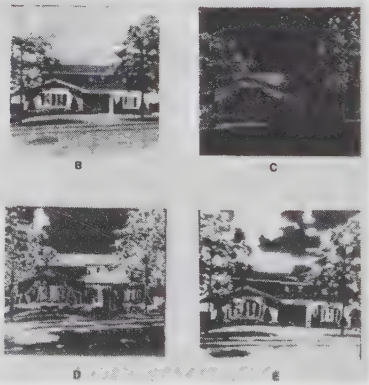


ABBILDUNG 7 :

FLÄCHENBILDUNG DURCH HISTOGRAMM-ANALYSEN

A - GRAUWERTHISTOGRAMM

B - GRAUWERTBILD (GRAUWERTE 0-255)

C - GRAUWERTSCHEIBE (32-120)

D - GRAUWERTSCHEIBE (121-183)

E - GRAUWERTSCHEIBE (184-255)

ABBILDUNG 8 :

FLÄCHENBILDUNG DURCH HISTOGRAMM-ANALYSEN

A - GRAUWERTHISTOGRAMM

B - GRAUWERTBILD (GRAUWERTE 0-255)

C - GRAUWERTSCHEIBE (48-119)

D - GRAUWERTSCHEIBE (120-207)

E - GRAUWERTSCHEIBE (208-243)

Die Tatsache, daß dieses Verfahren an diesem Beispiel zu guten Ergebnissen führt, täuscht über die Problematik des Ansatzes hinweg. Die Form des Histogrammes hängt kritisch vom Bildbereich ab, aus dem es ermittelt wird. Treten im Bildbereich viele, texturierte Regionen auf, so tragen Merkmale verschiedener Regionen häufig zu denselben Modalwerten bei oder führen - im schlimmsten Fall - dazu, daß keine Modalwerte mehr auftreten. Abbildung 8 weist ein solches Beispiel auf. Trotz der Vielzahl der verschiedenen Regionen (Abb. 8 B) treten im Histogramm (Abb. 8 A) nur zwei ausgeprägte Modalwerte auf. Demzufolge enthalten die hier ausgesuchten Grauwertscheiben mehrere verschiedene Regionen, wobei einige Regionen - z.B. die Bäume - in allen drei Grauwertscheiben (Abb. 8 C - E) auftreten.

Da verschiedene Regionen im Bild zu einem Wertebereich des Histogrammes beitragen, treten bei der Wahl von Schwellwerten Fehler auf, die dazu führen, daß gleichzeitig mehrere Regionen erfaßt werden oder Regionen fälschlicherweise in mehrere Teilregionen zerlegt werden. Bessere Ergebnisse sind nur zu erzielen, wenn die Histogramme nicht über dem gesamten Bild, sondern über Bildausschnitten berechnet werden. Dann allerdings werden Lage und Größe der Bildausschnitte zu kritischen Parametern, und es treten dieselben Probleme auf wie beim Einsatz von lokalen Operatoren der Kantendetektion.

KRITIK

Um eine Kritik der in Abschnitt 3.1 und 3.2 aufgeführten Verfahren von einem einheitlichen Standpunkt aus zu ermöglichen, ist es notwendig, einen solchen Standpunkt in Form eines - hypothetischen - Verarbeitungsschemas zu formulieren. Ein geeignetes Schema wird in dieser Arbeit als Folge von Übergängen von Repräsentationen postuliert. Repräsentationen sind durch die Art ihrer Daten und ihre Datenstrukturen gekennzeichnet. Die hier postulierte Folge baut auf drei Typen von Repräsentationen auf:
(1) Primär-Repräsentation (P-Repräsentation): diese Repräsentation ist der Ausgangspunkt der Szenenanalyse. Die Wahl der P-Repräsentation legt fest, welche Techniken niedererer Prozesse anwendbar sind. Als Daten stehen (a) die ursprünglichen Grauwerte oder (b) aus ihnen abgeleitete Daten zur Verfügung. Beispiele für die Erzeugung abgeleiteter Daten sind: die Anwendung lokaler Operatoren für Kompression und Kantendetektion, Fourier-Transformationen, numerische Approximationen oder die Auswertung

von Grauwertübergängen. Als Datenstrukturen liegen vor: (a) die ursprüngliche Grauwertmatrix, (b) Pyramidensysteme und (c) Baumstrukturen. Bei der Auswertung der P-Repräsentation werden Daten zusammengefaßt, die "zusammengehören". Bei Konturverfahren ist dies die Aufgabe der Linienverfolgung, bei Verfahren, die Regionen erfassen, die Aufgabe der Flächenbildung.

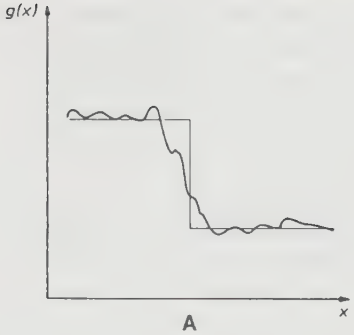
(2) Sekundär-Repräsentation (S-Repräsentation): in dieser Repräsentation werden die Ergebnisse der Auswertung der P-Repräsentation gespeichert. Die Auswertung der P-Repräsentation nach verschiedenen Kriterien kann zu widersprüchlichen Ergebnissen führen, die in der S-Repräsentation bereinigt werden. Die S-Repräsentation enthält Konturen und Regionen, denen Merkmale zugeordnet werden, die Information über Länge, Richtungen, Winkel bzw. Größe, Formen, Texturen, usw. tragen.

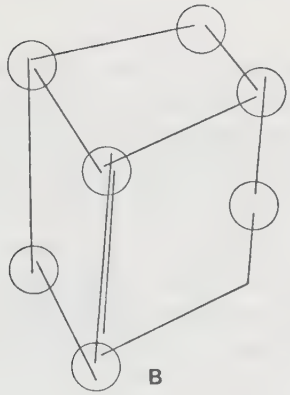
(3) Terminal-Repräsentation (T-Repräsentation): diese Repräsentation liefert eine Beschreibung der Szene, indem Regionen im Bild internen Modellen zugeordnet werden. Für die Zuordnung stehen die Informationen zur Verfügung, die in der S-Repräsentation vorliegen. Damit stellt die S-Repräsentation die Schnittstelle zwischen niederen und höheren Prozessen dar; die T-Repräsentation liefert das Ergebnis der höheren Prozesse.

Obwohl nur wenige Ansätze explizit diese drei Typen von Repräsentationen enthalten (z.B. /32, 45/), erlaubt dieses Schema eine Beurteilung einer Reihe von Verfahren. Im folgenden werden einige wichtige Kritierien für die Beurteilung von Verfahren niederer Prozesse erörtert.

DIE WAHL DER GRUNDELEMENTE VON P-REPRÄSENTATIONEN

Als Grundelemente von P-Repräsentationen stehen die ursprünglichen Grauwerte oder abgeleitete Daten zur Verfügung. Grauwerte einzelner Bildpunkte bieten nur wenig spezifische Information für die weitere Verarbeitung, solange nicht der Kontext jedes Bildpunktes berücksichtigt wird. Die Grauwertmatrix ist in diesem Fall eine triviale P-Repräsentation, zur Verarbeitung können nur einfache Techniken der Flächenbildung angewandt





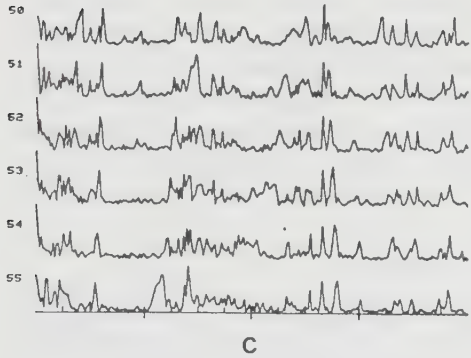


ABBILDUNG 9 : PROBLEME DER KANTENDETEKTION UND LINIENVERFOLGUNG

A - MODELL EINER KANTE (DÜNN) UND TYPISCHER VERLAUF EINER KANTE (DICK)IN EINEM GRAUWERTBILD

B - PROBLEMZONEN (KREISE), DIE SICH ALS FOLGE DER KANTENDETEKTION ERGEBEN : UNTERBRECHUNGEN, DOPPELTE LINIEN, ZU KURZE ODER ZU LANGE LINIEN, FALSCHE STEIGUNGEN, LINIEN SCHNEIDEN SICH NICHT IN EINER ECKE,.....

C - BILDZEILEN EINES GRADIENTENBILDES : DIE LINIENVERFOLGUNG MUSS AUS DIESEN DATEN LINIEN EXTRAHIEREN

werden: Schwellwertverfahren mit Hilfe von Histogrammanalysen oder Schmelzen von Punkten ähnlicher Grauwerte.

Beispiele abgeleiteter Daten sind Kantenelemente und Texturelemente. Bei der Wahl von Kantenelementen läßt sich ein Modell dessen angeben, was ein solches Modell im Idealfall ist - z.B. eine Grauwertstufe. Abbildung 9 zeigt, welche praktischen Probleme mit der Kantendetektion verbunden sind. Kanten treten in den seltensten Fällen als steile Grauwertstufe auf, son-

dern sind fast immer gestört und ziehen sich über mehrere Bildpunkte hinweg (Abb. 9 A). Die Folgen sind: es werden zuviele Bildpunkte als zu Kanten gehörig ermittelt, so daß zu dicke oder doppelte Linien entstehen; umgekehrt werden nicht alle Kantenpunkte detektiert, so daß sich unterbrochene Linien ergeben (Abb. 98).

Es ist außerordentlich schwierig, Linien aus dem Gradientenbild zu extrahieren (Abb. 9 C verdeutlicht das Problem der Linienverfolgung). Fast immer können nur unvollständige und gestörte Ergebnisse in die S- Repräsentation als Liniendarstellung eingetragen werden.

Bei der Wahl von Texturelementen ist es aufgrund der Vielzahl möglicher Texturen wesentlich schwieriger, Modelle idealer Elemente anzugeben. Die meisten Ansätze sind auf die Texturen eines vorliegenden Bildes zugeschnitten und erfassen nur Aspekte der Texturen, die sich aus einfachen Auszählungen von Grauwertübergängen, Lauflängen, usw. ergeben. Die Folgen sind: Regionen werden nicht scharf abgegrenzt und fließen ineinander oder werden umgekehrt in viele kleine Regionen aufgespalten.

LOKALE ENTSCHEIDUNGEN ; DAS FENSTERPROBLEM

Beim Einsatz von Operatoren, die auf Fenstern definiert sind, tritt stets das bekannte "Fensterproblem" auf. Es besteht darin, daß Lage und Größe des Fensters kritisch in die Ergebnisse der Operatoren eingehen. Das Fensterproblem tritt auf fast allen Verarbeitungsebenen auf. Fenster werden bei der Extraktion lokaler Merkmale (z.B. Kantenelemente), der Ermittlung globaler Information (z.B. Histogramme) oder der Konstruktion hierarchischer Repräsentationen (bei Pyramidensystemen und Baumstrukturen) eingesetzt. Die Größe des Fensters muß so gewählt werden, daß das gesuchte Element (Kante, Region,...) möglichst gut überdeckt wird. Zu kleine Fenster führen dazu, daß lokale Störungen zu falschen Ergebnissen führen. Zu große Fenster verwischen Details und führen ebenfalls zu falschen Ergenissen. Neben der Größe ist auch die Plazierung der Fenster ein kritischer Parameter. Fenster, die auf Grenzlinien von Regionen liegen, erfassen Bildpunkte beider Regionen und führen zu falschen Werten.

DIE EXTRAKTION ABGELEITETER DATEN FÜR P-REPRÄSENTATIONEN

Abgeleitete Daten entstehen häufig durch den Vergleich der Grauwertverteilung in einem Fenster mit einem idealen Modell eines Grundelementes, indem eine Faltung der Operatorfunktion mit der Grauwertfunktion vorgenommen wird. Es kann nicht erwartet werden, daß ein einziges Modell eines Grundelementes für die Analyse natürlicher Szenen ausreicht. Statt dessen ist ein Vielzahl von Grundelementen notwendig, die z.B. verschiedene Typen von Kantenelementen berücksichtigen. Zusätzlich ist es notwendig, auch mit Modellen verschiedener Größe zu arbeiten. Beim Übergang von der Grauwertmatrix zu einer P-Repräsentation mit abgeleiteten Daten muß also die Grauwertverteilung mit vielen Modellen verschiedenen Typs und verschiedener Größe verglichen werden. Aus den Ergebnissen dieses Vergleiches muß dasjenige Modell ausgesucht werden, welches "am besten paßt" /44/. Diese Entscheidung ist sehr schwierig und häufig fehlerbehaftet, da zu diesem Zeitpunkt der Verarbeitung nur lokale Information vorliegt. Diese Fehler beeinflussen die gesamte nachfolgende Verarbeitung und lassen sich kaum rückgängig machen.

INFORMATIONSFLUß

Vielen Ansätzen liegt ein Informationsfluß zugrunde, bei dem jede Verarbeitungsstufe nur die Information der vorangegangenen Verarbeitungsstufe zur Verfügung hat. Der Nachteil dieser Vorgehensweise liegt darin, daß jede Stufe auf einer Kontext-Ebene arbeitet, in der jeweils lokale Information der vorherigen Kontext-Ebene vorliegt; so werden z.B. auf der Ebene der Bildpunkte nur Punkte innerhalb einer kleinen Umgebung zur Ermittlung von Grauwertänderungen herangezogen, auf der Ebene der Linienverfolgung werden nur nahe beieinanderliegende Kanten- oder Linien-Elemente berücksichtigt, und so weiter. Da jede Verarbeitungsstufe aus einer Kontext-Ebene die nächsthöhere erarbeitet, ist es zu keiner Zeit möglich, gewissermaßen "nach oben" auf eine höhere Kontext-Ebene zu schauen.

3.3 EIN KATALOG VON VERARBEITUNGS-PRINZIPIEN

Ziel dieses Abschnittes ist es, die Erfahrungen, die für eine Reihe von Verfahren der Szenenanalyse vorliegen, in Verarbeitungsprinzipien umzusetzen, die als Richtlinien für die Entwicklung neuer Ansätze dienen. Zunächst sind einige Definitionen notwendig:

Definition: Eine Kontext-Ebene ist gegeben durch eine Menge K von Elementen und ein Abstandsmaß d zwischen diesen Elementen:

(1) $K := \{{}^{j}E_i / j=1,\ldots,N \wedge i=1,\ldots,T\}$.

Dabei bezeichnet Index j ein spezielles Element und Index i die Element-Klasse, zu der das Element gehört.

(2) $d({}^{j_1}E_i, {}^{j_2}E_i)$ legt den Abstand zwischen zwei Elementen fest.

Die Element-Klassen enthalten als Elemente beispielsweise Bildpunkte, Kanten-, Linien-, Textur-Elemente, Regionen, zusammengesetzte Regionen, usw. Zwischen den Element-Klassen wird eine (Halb-)-Ordnung vorausgesetzt, die eine Hierarchie der Element-Klassen bezüglich der Komplexität der Elemente definiert. Eine Formalisierung dieser Zusammenhänge wird hier nicht angestrebt. Statt dessen wird der Gedanke der Hierarchie der Element-Klassen an einem Beispiel verdeutlicht: die unterste Ebene bildet die Klasse der Bildpunkte; die nächsthöheren Ebenen bilden die Klassen der Kanten-, Linien- oder Texturelemente; auf diese folgen Ebenen mit Klassen von Regionen und so fort. Als Abstandmaß bietet sich der euklidische Abstand an, es sind jedoch auch andere Maße möglich (z.B. die sogenannte Manhattan-Metrik, bei der der Abstand zweier Punkte gleich der Weglänge in einem rechtwinkligen Straßen-Netz ist).

Definition: Als Kontext eines Elementes ${}^{j}E_i$ innerhalb einer Kontext-Ebene gilt die Menge der Elemente $\{{}^{k_n}E_i\}$, für die gilt:

(1) $d({}^{k_n}E_i, {}^{j}E_i) < \theta$

(2) ein vorgegebenes Prädikat P trifft auf alle ${}^{k_n}E_i$ zu.

Durch (1) werden die "benachbarten Elemente" eines Elementes bestimmt, durch (2) diejenigen ausgewählt, die bei der Verarbeitung eines Elementes berücksichtigt werden. Dies können in einfachen Fällen z.B. alle Elemente sein, die innerhalb eines Fensters um Element ${}^{j}E_{i}$ liegen. Das Prädikat P kann jedoch auch Eigenschaften der benachbarten Elemente berücksichtigen, indem z.B. benachbarte Elemente gleicher Intensität den Kontext bilden. Verfahren, bei denen nicht nur die Lage von benachbarten Elementen, sondern auch deren Eigenschaften erfaßt werden, sollen im folgenden als kontext-abhängig bezeichnet werden.

Eine Abgrenzung der Begriffe lokal/global muß berücksichtigen, daß es sich um zweideutige Begriffe handelt: zum einen können diese Begriffe innerhalb einer Kontext-Ebene angewandt werden, zum anderen (wegen der Hierarchie der Element-Klassen) zwischen Kontext-Ebenen. Innerhalb einer Kontext-Ebene sind lokale Verfahren dadurch gekennzeichnet, daß sie nur wenige benachbarte Elemente erfassen, während bei globalen Verfahren mehr Kontext berücksichtigt wird. Dabei ist zu berücksichtigen, daß es nicht immer sinnvoll ist, globale Information durch Berücksichtigung von mehr Kontext zu gewinnen: das den Kontext definierende Prädikat kann meistens nur allgemein vorgegeben werden, wenn es auf viele benachbarte Elemente zutreffen soll. Statt dessen kann globale Information häufig durch den Übergang zu einer höheren Kontext-Ebene erreicht werden. Innerhalb einer Kontext-Ebene sind also die Begriffe lokal/global durch den Umfang der Kontextes abgegrenzt, zwischen Kontext-Ebenen ist die Abgrenzung durch die Hierarchie der Element-Klassen gegeben.

Auf diesen Definitionen und Abgrenzungen aufbauend, können einige Prinzipien für die Entwicklung von Verfahren formuliert werden:

DAS PRINZIP DER MEHRFACHEN REPRÄSENTATIONEN

Dieses Prinzip wurde bereits in Abschnitt 3.2 als hypothetisches Verarbeitungsschema entwickelt. Ihm liegt der Gedanke zugrunde, den Vorgang der Szenenanalyse als Folge von Übergängen verschiedener Repräsentationen aufzufassen. Diese Folge baut auf drei Typen von Repräsentationen auf: Primär-, Sekundär- und Terminal-Repräsentation. Diese haben unterschiedliche Aufgaben und setzen sich aus unterschiedlichen Elementen zusammen

(vgl. Abschnitt 3.2). Die P-Repräsentation muß eine weitgehend exakte Beschreibung der Grauwertmatrix darstellen. Die Elemente der P-Repräsentation enthalten Kontext-Information, die Struktur der P-Repräsentation ist hierarchisch, um globale und lokale Analysen zu unterstützen. Die Verarbeitung der P-Repräsentation muß nach mehreren Gesichtspunkten erfolgen, wobei Kantendetektion und Linienanalysen sowie Flächenbildung und Texturanalysen besonders wichtig sind. Die Verarbeitung der P-Repräsentation führt zur S-Repräsentation, in der die Ergebnisse der verschiedenen Auswertungen zusammengetragen werden. Hier werden widersprüchliche Ergebnisse bereinigt, bis konsistente Ergebnisse vorliegen. Die S-Repräsentation enthält Konturen und Regionen, aus denen Merkmale für Formen, Texturen, usw. berechnet werden. Mit Hilfe dieser Merkmale wird eine Analyse der Konturen und Regionen vorgenommen, deren Ergebnis

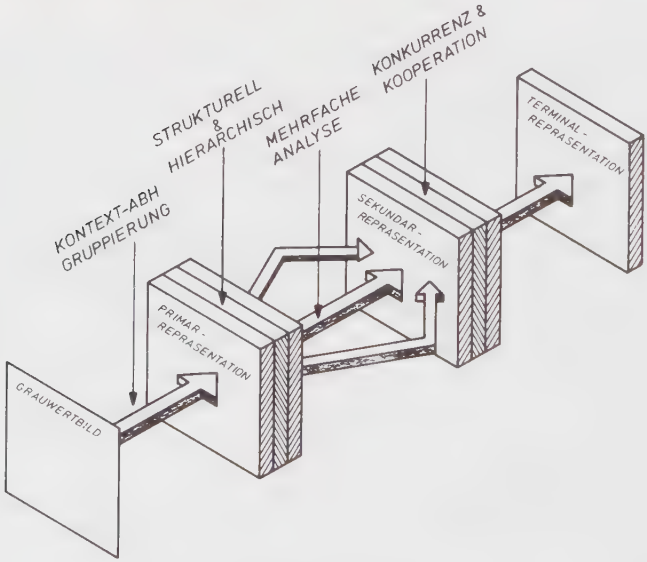


ABBILDUNG 1o : DER VORGANG DER SZENENANALYSE ALS EINE FOLGE VON ÜBERGÄNGEN VON REPRÄSENTATIONEN

in der T-Repräsentation festgehalten wird. In der T-Repräsentation liegt eine Beschreibung der Szene durch interne Modelle vor. Abbildung 10 verdeutlicht den Zusammenhang zwischen den verschiedenen Repräsentationen und zeigt, wo einige der folgenden Prinzipien eingreifen.

DAS PRINZIP DER STRUKTURELLEN P-REPRÄSENTATION

Das Fensterproblem und das Problem des Vergleichs der Grauwertverteilung mit idealen Modellen zur Extraktion von Grundelementen führen zu Fehlern bei der Erzeugung der P-Repräsentation, weil zu diesem Zeitpunkt nur lokale Information zur Verfügung steht. Zur Vermeidung solcher Fehler ist es notwendig, die Struktur des ursprünglichen Grauwertverlaufes zu erhalten und eine möglichst exakte Beschreibung der Struktur als P-Repräsentation zu verwenden. Statt Auswertungen und Vergleiche vorzunehmen, wird auf diese Weise lediglich Kontext-Information aufbereitet, die die nachfolgenden Auswertungen unterstützt.

DAS PRINZIP DER KONTEXT-ABHÄNGIGEN GRUPPIERUNG

Strukturen von Grauwertverteilungen sind durch Relationen benachbarter Bildpunkte definiert und müssen durch kontext-abhängige Kriterien detektiert werden. Durch Gruppierung benachbarter Bildpunkte, die den Kriterien entsprechend "zusammengehören", werden Teilstrukturen erfaßt und beschrieben. Durch die Formulierung kontext-abhängiger Prinzipien wird das Fensterproblem vermieden.

DAS PRINZIP DER HIERARCHISCHEN P-REPRÄSENTATION

Zur Vermeidung einer Verarbeitung, die ausschließlich auf lokaler Information aufbaut, ist es notwendig, daß die Verarbeitungsmoduln auf verschiedenen Kontext-Ebenen miteinander kommunizieren können. Dies erfordert eine hierarchische P-Repräsentation. Durch schnelle Verarbeitung grober Daten wird eine globale Analyse durchgeführt, die Hypothesen über Konturen und Regionen liefert. Diese vorläufigen Ergebnisse leiten die sorgfältige Analyse, indem sie Information über Grobstrukturen bereitstellen.

DAS PRINZIP DER MEHRFACHEN AUSWERTUNG

Verfahren, die allein auf Konturen bzw. Regionen aufbauen, liefern beim heutigen Stand der Technik nur unvollständige Ergebnisse. Deshalb muß die P-Repräsentation mehrfach ausgewertet werden, um sowohl Information über Konturen als auch über Regionen un die S-Repräsentation eintragen zu können.

DAS PRINZIP DER KONKURRENZ UND KOOPERATION

Die Ergebnisse, die durch die mehrfache Auswertung der P-Repräsentation vorliegen, müssen miteinander verglichen und abgestimmt werden. An einigen Orten können Ergebnisse sich widersprechen, während sie sich an anderen Orten gegenseitig unterstützen. Der Grundgedanke der Konkurrenz und Kooperation besteht darin, daß sich die Ergebnisse gegenseitig beeinflussen, abschwächen oder verstärken, bis sich ein konsistentes Ergebnis eingestellt hat /2/. (Allerdings ist damit noch nicht gewährleistet, daß das Ergebnis auch das "Richtige" ist.)

DAS PRINZIP DER ZÖGERNDEN FALSIFIKATION

Dieses Prinzip betrifft die Verarbeitungsgeschwindigkeit, mit der von einer Repräsentation zur nächsten übergegangen wird. Es entspricht ungefähr dem "Principle of Least Commitment" /44/ und besagt, daß Übergänge so vorsichtig wie möglich vollzogen werden sollten. Zu Beginn einer Verarbeitung liegt eine Vielzahl möglicher Alternativen vor. Entscheidungen über die Auswahl von Alternativen sollten erst dann getroffen werden, wenn genügend Information für eine zuverlässige Entscheidung vorliegt. In der Regel ergeben sich auf jeder Verarbeitungsstufe Widersprüche, die verschiedene Alternativen ausschließen. Durch Falsifikation von Hypothesen wird die Menge der möglichen Alternativen ständig verringert bis ein endgültiges Ergebnis vorliegt. Falsifikationen ergeben sich aus der Verarbeitung des Kontextes, in den die Alternativen hineinpassen müssen.

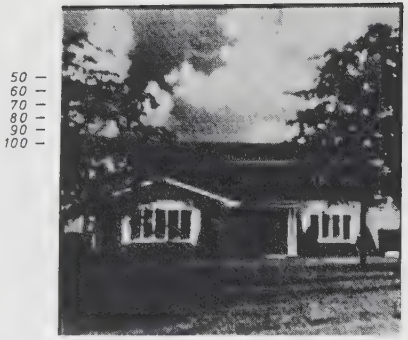
Zur Verwirklichung dieser Prinzipien sind einige Voraussetzungen notwendig. Es müssen große Datenmengen gespeichert werden; die mehrfache Verarbeitung setzt hohe Verarbeitungsgeschwindigkeit voraus; Informa-

tionen müssen zwischen verschiedenen Moduln hin und her fließen; eine Explosion des Suchraumes bei der Auswahl von Alternativen muß vermieden werden; es müssen Maße entwickelt werden, die ausdrücken, wie gut eine Analyse ist (die Konfidenz einer Analyse) ; Methoden für die Repräsentation von Wissen und internen Modellen müssen entwickelt werden. Einige dieser Voraussetzungen sind beim heutigen Stand der Technik nur bedingt erfüllt. Die vorliegende Arbeit leistet im folgenden Beiträge, die diese Prinzipien berücksichtigen. Probleme der strukturellen, hierarchischen P-Repräsentation und der kontext-abhängigen Gruppierung werden in Kapitel 4. behandelt; Kapitel 5. befaßt sich mit Problemen der mehrfachen Auswertung, wobei ebenfalls Probleme der kontext-abhängigen Gruppierung untersucht werden. In beiden Kapiteln werden hierzu Algorithmen entwickelt, die implementiert und experimentell untersucht sind.Beiträge zur mehrfachen Repräsentation und zur Kooperation und Konkurrenz werden in Kapitel 6. in einem gedanklichen Modell entwickelt.

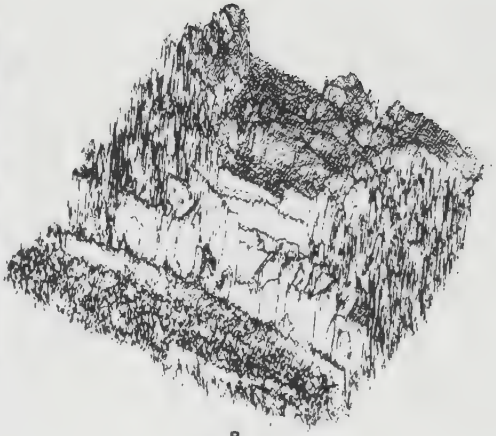
4. RELATIONAL-BÄUME: EIN NEUER ANSATZ ZUR REPRÄSENTATION VON GRAUWERTBILDERN

Der vorliegende Ansatz baut auf folgender Grundidee auf: gemäß den in Abschnitt 3.3 entwickelten Prinzipien wird die diskrete Grauwertmatrix $G(x_i, y_j)$ in eine Datenstruktur überführt, bevor Auswertungen vorgenommen werden. Diese Datenstruktur stellt eine symbolische Beschreibung der Struktur des Grauwertverlaufes im Bild dar. Aus dieser Struktur werden anschließend Teilstrukturen extrahiert. Diese Auswertungen wurden durch Kontext-Information unterstützt, die in der hierarchischen Datenstruktur zur Verfügung steht.

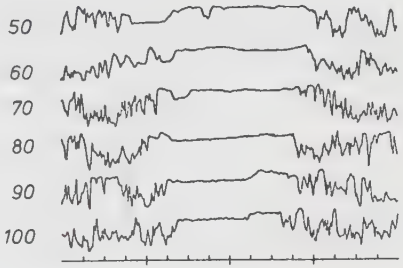
Ein wesentlicher Gedanke des Konzepts ist es, die Grauwertmatrix als Grauwertgebirge aufzufassen. Abbildung 11 A zeigt ein Grauwertbild einer natürlichen Szene, Abbildung 11 B das entsprechende Grauwertgebirge. Die exakte Erfassung der dreidimensionalen Struktur des Grauwertgebirges ist äußerst schwierig und mit sehr hohem Arbeitsaufwand verbunden. In technischen Systemen werden Bilder häufig zeilenweise abgetastet und verarbeitet. Es liegt daher nahe, nicht die Grauwert - matrix $G(x, y)$ in eine hierarchische Datenstruktur zu überführen, sondern statt dessen jede Bildzeile einzeln durch eine hierarchische Datenstruktur darzustellen. An die Stelle des Grauwertgebirges tritt eine Folge von Intensitäts-Profilen, deren Struktur mit geringerem Aufwand erfaßt werden kann. Abbildung 11 C zeigt einige Intensitäts-Profile (I-Profile) des Grauwertbildes aus Abb. 11 A. Die I-Profile werden im folgenden durch $I_j(x_i)$ bezeichnet (auf den Zeilenindex j kann im folgenden verzichtet werden). Die zeilenweise Verarbeitung ist besonders für die heutigen, sequentiell arbeitenden Computer geeignet; Nachteile der zeilenweisen Verarbeitung müssen durch eine zusätzliche Verarbeitung quer zu den Zeilen ausgeglichen werden.



A



B



C

ABBILDUNG 11 : BETRACHTUNGSWEISEN EINES GRAUWERTBILDES

A - GRAUWERTBILD (256 x 256 BILDPUNKTE)

B - GRAUWERTGEBIRGE (DARSTELLUNG UM 30^{0} NACH UNTEN GEDREHT)

C - FOLGE VON INTENSITÄTSPROFILEN (DIE ENTSPRECHENDEN BILDZEILEN SIND IN ABB.11 A ANGEZEICHNET)

4.1 DIE REPRÄSENTATION VON BILDZEILEN DURCH RELATIONAL-BÄUME

Ein Vergleich des Grauwertbildes in Abbildung 11 A mit dem Grauwertgebirge (Abbildung 11 B) zeigt, daß die texturierten Regionen im Bild (z.B. die Bäume oder das Gras) durch dicht verteilte Maxima und Minima im Grauwertgebirge gekennzeichnet sind. Wenig texturierte oder homogene Regionen hingegen (z.B. das Haus oder der Himmel) liegen im Grauwertgebirge als ebene Flächen vor. Oberträgt man diese Beobachtungen auf I-Profile, so stellt man fest: texturierte Regionen werden in I-Profilen durch Relationen von Maxima und Minima charakterisiert, während homogene Regionen in I-Profilen als Plateaus auftreten (vgl. Abb. 11 C).

In diesem Abschnitt wird zunächst ein von EHRICH & FOITH entworfenes Konzept für einen Relational-Baum (R-Baum) entwickelt, der die Beschreibung texturierter Regionen ermöglicht /17, 19/. In Abschnitt 4.2 wird die Datenstruktur verallgemeinert, um auch homogene Regionen beschreiben zu können.

Die Betrachtung einer Bildzeile als I-Profil hat folgende Konsequenzen: es werden weder die Bildpunkte x betrachtet noch deren Grauwerte I(x), sondern "Profilpunkte" mit den Koordinaten (x,I(x)). Die Relationen zwischen Profilpunkten werden ausgewertet, um die Struktur des I-Profils zu erhalten. Aussagen über die Zusammengehörigkeit von Bildpunkten ergeben sich durch Projektion zusammengehöriger Profilpunkte auf die X-Achse. Die folgenden Definitionen beziehen sich auf Profilpunkte; Aussagen über Bildpunkte werden erst bei der Auswertung der R-Bäume wichtig und ergeben sich durch die Projektion entsprechender Profilpunkte auf die X-Koordinaten in einfacher Weise.

Intuitiv kann ein I-Profil strukturell beschrieben werden, indem man es in Segmente zerlegt und die Segmente zusammen mit ihren gegenseitigen Relationen angibt. Die Zerlegung darf nicht von der Vorgabe von Fenstern oder anderen Parametern abhängen. Statt dessen werden Segmente durch eine kontext-abhängige Gruppierung benachbarter Bildpunkte gewonnen. Zur Erfassung von Texturen wird die Relation "Berg" als Gruppierungskriterium gewählt /17/.

Ein "Berg" besteht aus Profilpunkten einer lokalen Nachbarschaft δ von Profilpunkten, die sich um ein lokales Maximum bezüglich d er Grauwerte scharen und oberhalb des höchsten benachbarten Minimums liegen. Abbildung 12 A zeigt ein I-Profil und seine Zerlegung in "Berge": wir betrachten zunächst ein lokales Minimum innerhalb einer Umgebung δ als ein "Tal" T mit den Koordinaten $(x^T, I(x^T))$:

Definition: Ein Tal T ist gegeben durch:

$$(x^T, I(x^T)) := \min_I \{ (x, I(x)) \ / \ x \in [x^T - \frac{\delta}{2}, x^T + \frac{\delta}{2}] \} \qquad (4.1)$$

Zieht man nun einen waagrechten Strich $y = I(x^T)$ rechts und links von T, so ergeben sich an den Profilpunkten $(x, I(x))$, an denen zum ersten Mal wieder $I(x) = I(x^T))$ ist, ein rechter Basispunkt R und ein linker Basispunkt L, für deren Koordinaten gilt:

$$L := (x^L, I(x^L)) \text{ mit } x^L = \max_x \{ x \ / \ I(x) \leq I(x^T) \wedge x < x^T \} \qquad (4.2)$$

bzw.

$$R := (x^R, I(x^R)) \text{ mit } x^R = \min_x \{ x \ / \ I(x) \leq I(x^T) \wedge x > x^T \} \qquad (4.3)$$

Eine Basis ist die Menge der Punkte des waagrechten Striches zwischen L und T bzw. T und R:

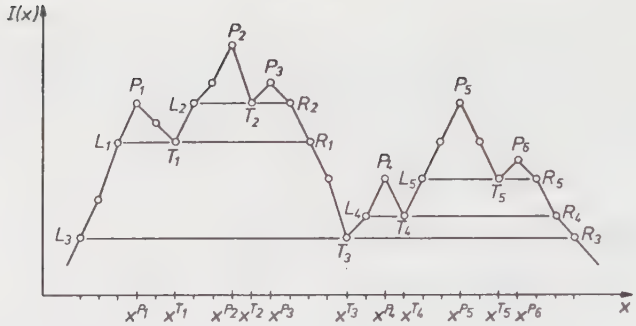
Definition: Eine Basis zwischen L und T bzw. T und R ist: (4.4)

$$\text{BASIS} := \{ (x_i, I(x^T)) / \ x^L \leq \ldots \leq x_i \leq \ldots \leq x^T$$
$$\text{bzw. } x^T \leq \ldots \leq x_i \leq \ldots \leq x^R \}$$

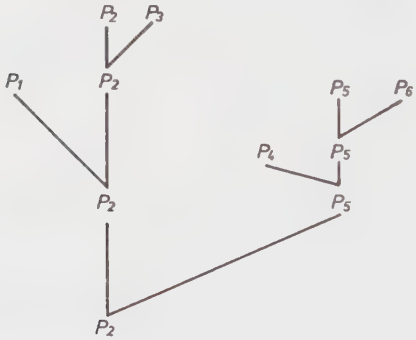
Ein Berg ist die Menge der Profilpunkte, die oberhalb einer Basis zwischen L und T bzw. T und R liegen:

Definition:

$$\text{BERG} := \{ (x_i, I(x_i)) \ / \ I(x_i) \geq I(x^T) \text{ für alle } x_i \text{ mit:} \qquad (4.5)$$
$$x^L \leq \ldots \leq x_i \leq \ldots \leq x^T \text{ bzw. } x^T \leq \ldots \leq x_i \leq \ldots \leq x^R \}$$



A Die Zerlegung eines I-Profiles in Berge



B Darstellung der Dominanz von Bergen durch einen Relational - Baum

ABBILDUNG 12 : DIE BESCHREIBUNG EINES INTENSITÄTS-PROFILES DURCH EINEN RELATIONAL-BAUM

Berge sind die Grundelemente der P-Repräsentation. Sie werden im folgenden durch ihre höchsten Punkte P_i gekennzeichnet. Beim Betrachten der I-Profile in Abbildung 12 A fällt auf, daß Berge ineinander verschachtelt sind, wobei bestimmte Berge andere dominieren. Im Beispiel (vgl. Abbildung 12) dominiert P_2 oberhalb der Basis zwischen L_3 und T_3 die Berge P_1 und P_3. Eine Relation "Dominanz" $D(P_i,P_j)$ drückt aus, daß P_i dominant für P_j ist, wenn beide Berge oberhalb einer Basis liegen und $I(x^{P_i}) \geq I(x^{P_j})$ ist. Die Dominanz erzeugt eine Ordnung über den Bergen und führt im Bei-

spiel von Abbildung 12 zu folgendem Ausdruck:

$$D(D(D(P_2,P_3),P_1),D(D(P_5,P_6),P_4). \qquad (4.6)$$

Dieser Klammerausdruck läßt sich unter Berücksichtigung der Reihenfolge der Berge von links nach rechts im I-Profil durch eine Baumstruktur darstellen, wenn zusätzlich verabredet wird, den dominanten Berg innerhalb einer Klammer als den Repräsentanten dieser Klammer einzusetzen (Abbildung 12 B). Für zwei Sonderfälle müssen Konventionen eingeführt werden (vgl. Abbildung 13):

(1) Falls N direkt benachbarte Minima die gleiche Intensität haben, werden diese N Minima durch einen Knoten repräsentiert (Abbildung 13A).

(2) Falls N direkt benachbarte Berge Maxima gleicher Intensität haben, wird ein N-Tupel dieser Berge als Repräsentant eingesetzt (Abbildung 13 B).

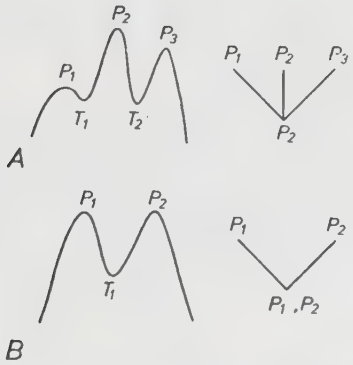


ABBILDUNG 13 : KONVENTIONEN FÜR DAS AUFTRETEN MEHRERER BENACHBARTER EXTREMA

A - MINIMA GLEICHER INTENSITÄT B - MAXIMA GLEICHER INTENSITÄT

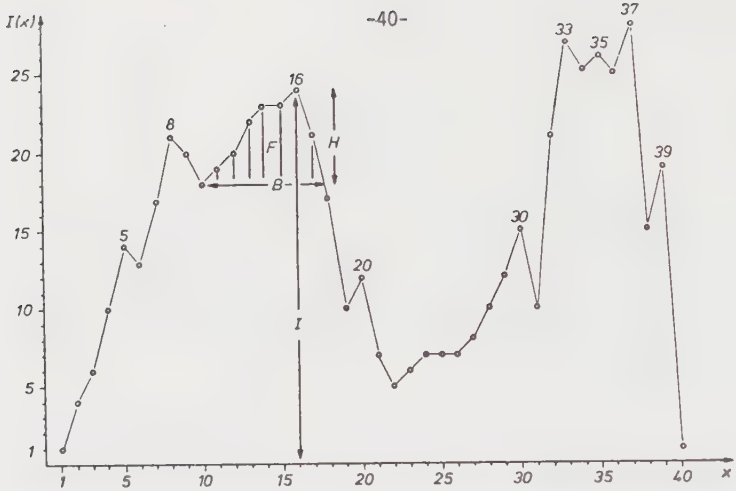


ABBILDUNG 14 : ATTRIBUTE VON BERGEN

I=ABSOLUTE HÖHE H=RELATIVE HÖHE B=BASISBREITE F=FLÄCHE

Der Relational-Baum (R-Baum) ist somit ein Graph, der die Verschachtelung von Bergen wiedergibt. R-Bäume werden mit obenliegenden Blättern dargestellt, da so der Zusammenhang zwischen Blättern und Bergen besser ersichtlich ist. Bezüglich der Bezeichnungen "Nachfolger" und "Vorgänger" gilt die übliche Sprechweise, d.h. die Wurzel hat keine Vorgänger, die Blätter haben keine Nachfolger.

Ein Berg hat eine Reihe von Attributen, die wichtige Informationen für die Auswertung darstellen und leicht berechnet werden können:

(A1) Die ausgezeichneten Koordinaten (vgl. Abb. 12 A):

x^L , x^P , x^T (falls $x^P < x^T$)

bzw. x^T , x^P , x^R (falls $x^P > x^T$)

(A2) Die absolute Höhe, d.h. $I(x^P)$

(A3) Die relative Höhe $H: = I(x^P) - I(x^T)$

(A4) Die Basisbreite $B: = x^T - x^L$ (falls $x^P < x^T$)

bzw. $B: = x^R - x^T$ (falls $x^P > x^T$)

(A5) Die Fläche F des Berges:

$$F: = \sum_{x=x^L}^{x^T} [I(x) - I(x^T)] \cdot \Delta x \quad (\text{falls } x^P < x^T)$$

bzw.

$$F: = \sum_{x=x^T}^{x^R} [I(x) - I(x^T)] \cdot \Delta x \quad (\text{falls } x^P > x^T)$$

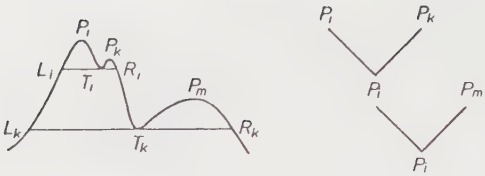
(Dabei ist Δx der Abstand zweier Bildpunkte und wird $\Delta x = 1$ gesetzt).

Abbildung 14 zeigt die Attribute A2 - A5 an einem Beispiel. Jedem Knoten des R-Baumes wird eine Liste dieser Attribute zugeordnet. Die Bezeichnung "Relational-Baum" wurde gewählt, da Berge durch die Relationen der lokalen Minima und Maxima definiert sind, und die Baumstruktur Relationen zwischen diesen Relationen wiedergibt.

Definition: Ein Relational-Baum ist ein Graph, dessen Struktur die Dominanz-Relation von Bergen wiedergibt.

Die Dominanz-Relation wird durch lokale Minima mit zunehmenden Intensitätswerten induziert, die in sich einschließenden Intervallen des I-Profiles liegen. Die Blätter des Baumes entsprechen Bergen, die keine weitere Struktur haben und sind durch diese Berge markiert. Alle anderen Knoten im Baum entsprechen zusammengesetzten Bergen und sind mit dem dominanten Berg markiert. Die Wurzel des Baumes entspricht dem gesamten I-Profil. Der Baum kann von den Blättern zur Wurzel oder umgekehrt konstruiert werden.

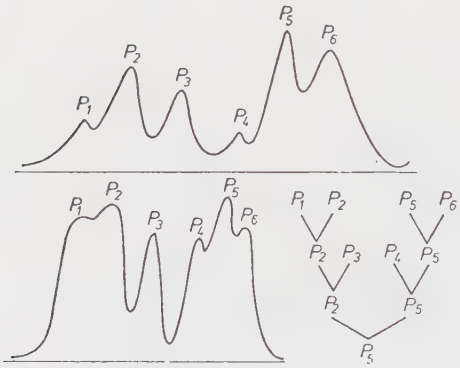
Ein R-Baum ist somit eine hierarchische Darstellung eines I-Profiles, in der durch mehrfache Gruppierungen kontext-abhängig immer größere Strukturen erfaßt werden. R-Bäume haben folgende Eigenschaften:



$x^{L_k} < x^{L_i} < x^{P_i} < x^{T_i} < x^{P_k} < x^{R_i} < x^{T_k} < x^{P_m} < x^{R_k}$

$I(x^{P_i}) > I(x^{P_k})$, $I(x^{P_i}) > I(x^{P_m})$, $I(x^{T_i}) > I(x^{T_k})$

A - durch R-Baum induzierte (Halb)-Ordnungen



B - Zwei Intensitätsprofile mit identischer Baumstruktur

ABBILDUNG 15 : ZUSAMMENHÄNGE DER STRUKTUR VON R-BÄUMEN MIT DER STRUKTUR VON I-PROFILEN

(E1) Ein R-Baum induziert eine totale Ordnung über den Bildpunkten und eine Halbordnung über den Grauwerten (Abbildung 15 A), d.h. die Struktur des I-Profiles bleibt erhalten.

(E2) Aufgrund von Eigenschaft E1 zerlegen R-Bäume die Menge der I-Profile in Äquvalenzklassen, deren Elemente sehr unterschiedlich in ihrer Form sein können (Abbildung 15 B).

(E3) Obwohl ein Berg mehrfach als Marke auftreten kann, hat die Marke bei jedem Auftreten eine andere Bedeutung: jedem Knoten ist eine "Attributs-Liste" zugeordnet, in der der entsprechende Berg durch die Attribute A1 - A4 gekennzeichnet ist. Für die Speicherung von A5 hat sich bisher keine Notwendigkeit ergeben. Bei Bedarf kann diese Liste durch weitere Attribute ergänzt werden.

(E4) Tritt ein Berg N-mal als Marke auf, so dominiert er N-1 andere Berge.

(E5) Wenn man in einem R-Baum von den Blättern in Richtung Wurzel auf einem Weg identischer Marken entlangläuft, nehmen die relativen Höhen und die Basisbreiten in der Attributsliste der durchlaufenen Knoten zu. Dadurch sind schnelle Suchverfahren möglich.

4.2 VERALLGEMEINERTE RELATIONAL-BÄUME

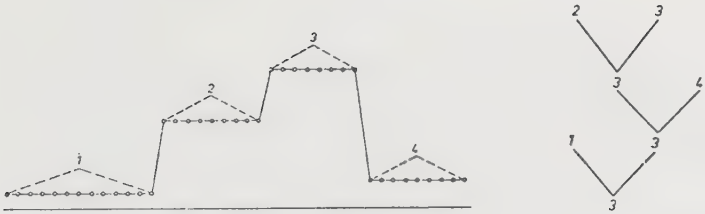
Durch die Wahl von Bergen als Grundelementen der P-Repräsentation werden zwei Arten von Profilpunkten bevorzugt, nämlich die eindimensionalen Minima und Maxima von I-Profilen. In Abbildung 11 wurde deutlich, daß die örtliche Verteilung von Minima und Maxima charakteristisch für Texturen ist. Diese Beobachtung wird durch Abbildung 16 unterstützt: Diese Abbildung zeigt die Lage aller x^P und x^T der R-Bäume, die aus den Bildzeilen des Grauwertbildes aus Abbildung 11A konstruiert wurden. Diese Punkte fallen im Bild in die texturierten Regionen. Auf diese Weise werden jedoch die homogenen Regionen des Bildes nicht erfaßt.

ABBILDUNG 16 : DIE ÖRTLICHE VERTEILUNG DER EINDIMENSIONALEN MINIMA UND MAXIMA IN ZEILEN-RICHTUNG IM GRAUWERTBILD AUS ABB. 11 A

Aus Abbildung 11 C kann der Grund hierfür abgelesen werden: diese Regionen treten in I-Profilen als Segmente mit Steigung Null auf. Solche Segmente sollen "Plateau" genannt werden.

Um auch Bilder natürlicher Szenen repräsentieren zu können, in denen sowohl Texturen als auch homogene Regionen auftreten, muß es möglich sein, in den R-Bäumen auch Plateaus zu erfassen und zu beschreiben. Dies geschieht, indem Plateaus als virtuelle Berge mit relativer Höhe Null aufgefaßt werden (siehe Abbildung 17). Durch die Einführung virtueller Berge über den Plateaus werden diese nahtlos in das Konzept des R-Baumes eingegliedert. Als absolute Höhe wird die Intensität des Plateaus gewählt, als Basisbreite die Länge des Plateaus. Je nach Lage des Plateaus wird an seinem rechten oder linken Ende ein virtuelles Tal angenommen, das einen Knoten im R-Baum erzeugt.

Definition: Ein verallgemeinerter Relational-Baum ist ein R-Baum, dessen Knoten Bergen oder Plateaus entsprechen.

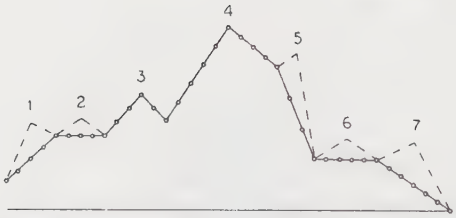


A - PLATEAUS ALS VIRTUELLE BERGE B - RELATIONAL-BAUM

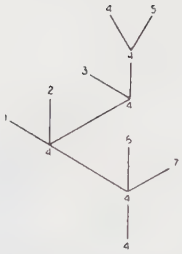
ABBILDUNG 17 : DIE BESCHREIBUNG VON PLATEAUS IN R-BÄUMEN

Grundsätzlich ist es möglich, die Zerlegung der I-Profile in Segmente zu verfeinern. So können als Grundelemente Segmente konstanter Steigung angenommen werden, die ebenfalls durch die Einführung virtueller Berge über den Segmenten in den R-Baum eingegliedert werden können (Abbildung 18). Daneben ist es auch möglich, die exakte Form der Flanken von Bergen zu erfassen, indem man diese Form in der Attributs-Liste der Knoten in einem Ketten-Kode /50/ speichert. Aus den bisherigen Ergebnissen von Experimenten ergibt sich keine Notwendigkeit für eine feinere Segmentation oder eine genauereErfassung der Form (vgl. Kapitel 5.).

An dieser Stelle ist ein Exkurs notwendig: digitalisierte Grauwertbilder sind in der Regel so verrauscht, daß Plateaus, wie die in Abb. 11 C gezeigten, nicht auftreten, da sie durch lokale Störungen in viele kleine Berge zerfallen. Um dennoch die Erfassung von Plateaus in R-Bäumen untersuchen zu können, wurden die Grauwertbilder deshalb durch eine Filterung im Ortsbereich geglättet. Hierfür wurde ein von EHRICH entwickeltes Verfahren eingesetzt /21/. Da Rauschunterdrückung nicht Thema dieser Arbeit ist, scheint diese Vorgehensweise gerechtfertigt. Die Filterung ist jedoch auch in R-Bäumen möglich, wobei gleichzeitig das Fensterproblem vermieden werden kann. Da die Struktur des I-Profiles im R-Baum repräsentiert ist, können gezielt solche Berge unterdrückt werden, die klein sind und vermutlich auf Rauschen zurückzuführen sind.



A - SEGMENTE KONSTANTER STEIGUNG



B - RELATIONAL-BAUM

ABBILDUNG 18 : DIE BESCHREIBUNG VON SEGMENTEN KONSTANTER STEIGUNG IN RELATIONAL-BÄUMEN

In Abschnitt 4.3 wird deutlich, daß sich eine Filterung auch während der Konstruktion der R-Bäume durchführen ließe: wenn bei der Baum-Konstruktion ein Knoten erzeugt wird, liegen gleichzeitig die Attribute A1 - A4 vor, aus denen entschieden werden kann, ob ein Berg eingetragen werden soll oder nicht. Sind die relative Höhe und die Basisbreite zu klein, so handelt es sich um Rauschen und es wird kein Knoten erzeugt. Liegen mehrere solcher Berge nebeneinander und ihre absoluten Höhen unterscheiden sich nur um einen kleinen Betrag, so werden diese Berge zu einem Plateau zusammengefaßt und in den R-Baum eingetragen. Allerdings besteht auch bei dieser Vorgehensweise das Problem , geeignete Parameter für die Attribute zu bestimmen. Eine solche Filterung wurde noch nicht implementiert, sie ist jedoch grundsätzlich möglich.

Weitere Modifikation der R-Bäume können auch dazu führen, das Konzept zu vereinfachen. So ist es möglich, auf die Baumstruktur zu verzichten und Berge und ihre Attribute in einer einfachen Liste zu speichern. In besonderen Fällen reicht dies z.B. für Textur-Analysen aus.

Das Konzept der R-Bäume ist speicheraufwendig (vgl. Abschnitt 4.3), erlaubt jedoch einfache Verarbeitungsverfahren (vgl. Kapitel 5.). Hoher Speicheraufwand zugunsten schneller Auswertungsverfahren ist beim Einsatz sequentiell arbeitender Computer eine wichtige Maßnahme, um Szenenanalysen mit einem vertretbaren Zeitaufwand durchführen zu können. Eine Anwendung des vorliegenden Konzeptes erscheint überall dort gerechtfertigt, wo Grauwertbilder natürlicher Szenen vorliegen - also z.B. bei Luftbild- oder Satellitenbildauswertungen.

4.3 DER ALGORITHMUS ZUR KONSTRUKTION VON RELATIONAL-BÄUMEN

Zur Konstruktion der R-Bäume wird ein schneller Keller-Algorithmus eingesetzt. Nach einer kurzen Darstellung der Grundidee des Algorithmus wird zunächst die ursprüngliche Version der Konstruktion von R-Bäumen beschrieben (vgl. Abschnitt 4.1); anschließend wird erörtert, wie verallgemeinerte R-Bäume (vgl. Abschnitt 4.2) erzeugt werden.

Der Algorithmus zur Konstruktion von R-Bäumen verwendet drei Keller, mit deren Hilfe kontext-abhängige Entscheidungen getroffen werden. Der Algorithmus durchläuft die Profilpunkte von links nach rechts (oder umgekehrt) und speichert :

(1) Profilpunkte im Breite-Keller ,

(2) Täler im Tal-Keller ,

(3) X-Koordinaten von Maxima zusammen mit Verweisen auf die Datenstruktur im Berg-Keller.

Entscheidungen werden stets durch Vergleiche der beiden oberen Eintragungen eines Kellers getroffen, so daß wenig gerechnet werden muß. Der Algorithmus befindet sich entweder im "Aufwärts"-Zustand (AUF) , d.h. $I(x_{i-1}) \leq I(x_i)$, oder im "Abwärts"-Zustand (AB), d.h. $I(x_{i-1}) \geq I(x_i)$.

Abbildung 19 zeigt ein einfaches Beispiel der Vorgehensweise: der Algorithmus durchläuft das I-Profil von links nach rechts. Im Aufwärts-Zustand werden Profilpunkte in den Breite-Keller eingetragen bis ein Maximum erreicht ist (Abb. 19 A). Dieses wird in den Berg-Keller eingetragen und der Abwärts-Zustand angenommen. Im Zustand AB werden die Eintragungen aller Profilpunkte, die oberhalb des gerade erreichten Profilpunktes $(x_i, I(x_i))$ liegen, aus dem Breite-Keller geleert - d.h. alle Punkte $(x_k, I(x_k))$, $k=1,\ldots,i-1$, für die gilt : $I(x_k) > I(x_i)$. Auf diese Weise liegen sich diejenigen Profilpunkte, deren X-Koordinaten die beiden oberen Eintragungen des Breite-Kellers sind, auf den Flanken eines Berges gegenüber, so daß sich bei Erreichen eines Minimums sofort die Basisbreite des Berges aus der Differenz der beiden oberen Keller-Elemente ergibt (Abb. 19 B). Nach Erreichen eines Minimums wird dieses in den Tal-Keller eingetragen und wieder der Aufwärts-Zustand angenommen (Abb. 19 C). Die Grundidee des Algorithmus beruht darauf, im Abwärts-

ABBILDUNG 19 : DIE KONSTRUKTION EINES R-BAUMES IN EINEM DURCHLAUF DURCH DIE PROFILPUNKTE VON LINKS NACH RECHTS

Zustand $I(x_i)$ mit $I(x^{T_n})$ desjenigen Minimums zu vergleichen, das zuletzt in den Tal-Keller eingetragen wurde. Ist $I(x_i) \leq I(x^{T_n})$, dann ist ein rechter Basispunkt R erreicht (vgl. Abb. 12 A), und es kann die Dominanz der beiden Berge entschieden werden, die links und rechts von T_n liegen. Die X-Koordinaten der betreffenden Berge sind die beiden oberen Eintragungen im Berg-Keller (Abb. 19 D). Nach Erzeugung eines Knotens im Baum, werden der nicht-dominante Berg und das Tal T_n aus dem Berg- bzw. Tal-

Keller entfernt (Abb. 19 E). Der Vorgang des Vergleiches von Keller-Eintragungen wird im Abwärts-Zustand fortgesetzt bis alle Berge und Täler abgearbeitet sind (Abb. 19 F). Mit Erreichen des letzten Profilpunktes liegt der komplette R-Baum vor.

Der R-Baum wird in einer mehrfach verketteten Liste, der Baum-Tabelle , gespeichert (Abbildung 2o). Die Tabelle besteht aus zwei Teil-Tabellen, der Tal-Tabelle und der Berg-Tabelle, in denen Information über Täler und Berge gespeichert wird. Zu jedem Tal in einem I-Profil gehören in der Regel zwei Berge, die links bzw. rechts vom Tal liegen. Für den Sonderfall benachbarter Täler gleicher Intensität (vgl. Abb. 13 A) wird nur das am weitesten links liegende dieser Täler in die Tal-Tabelle eingetragen; zu diesem Tal gehören dann auch die Berge der anderen gleichhohen Täler. Die Lage dieser Täler wird in der Attributs-Liste der Berge gespeichert.

Die Tal-Tabelle hat folgende Eintragungen :

LOC - Index der Eintragungen der Tal-Tabelle (Tal-Index)

DWN - Zeiger zum Index des Vorgängers, d.h. zum nächsten Knoten in Richtung Wurzel

TAL-INFO - die Lage des Tals : $(x^T, I(x^T))$

DOM - Zeiger zum Index des dominanten Berges, dessen X-Koordinate x^P die Marke des Knotens im Baum ist

PKA - Zeiger zum Index des Berges, der links vom Tal liegt.

Die Berg-Tabelle hat einen Eintrag für jeden Berg, bestehend aus :

ADR - Index der Eintragungen in der Berg-Tabelle (Berg-Index)

LCV - Zeiger zum Index des Tals, zu dem der betreffende Berg gehört

PK - X-Koordinate des Maximums : x^P

ATTRIBUTS LISTE - diese Liste enthält die Attribute A1 - A4 (vgl. Abschnitt 4.1)

NXT - Zeiger zum Berg-Index des rechten benachbarten Berges, der zum gleichen Tal gehört (Nachbar-Index).

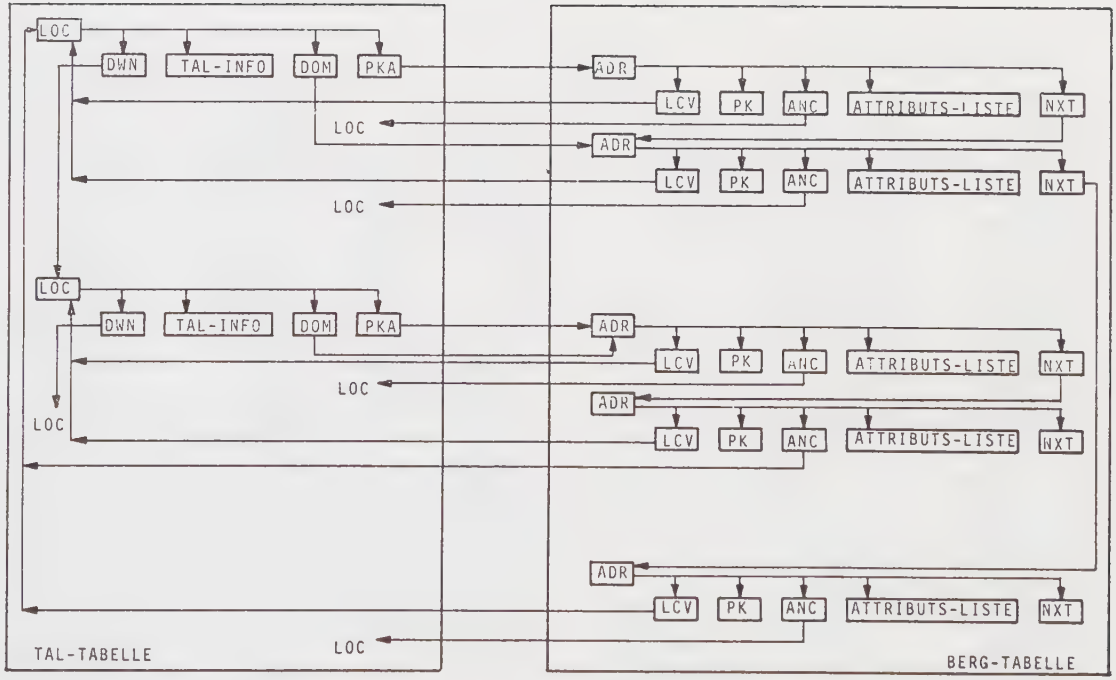


ABBILDUNG 2o : BAUM-TABELLE ZUR SPEICHERUNG EINES R-BAUMES

Der Aufwand der Speicherung ist durch die vielen Zeiger hoch. Diese erlauben jedoch schnelle Bewegungen im Baum. So ergibt sich z.B. ein Weg von einem Blatt zur Wurzel über eine Folge von Tal-Indizes (LOC); ein Blatt wird von der Wurzel durch eine Folge von Berg-Indizes (ANC) erreicht; benachbarte Berge eines Tals erhält man durch eine Folge von Nachbar-Indizes (NXT). Durch Kombinationen dieser Folgen kann jeder Knoten von jedem beliebigen Knoten aus ereicht werden. Dies ist für die Gewinnung von Kontext-Information wichtig.

Im folgenden wird eine detaillierte Beschreibung des Algorithmus zur Konstruktion eines R-Baumes gegeben. Dies erfordert eine Reihe von Definitionen, wobei $(x_i, I(x_i))$ stets den gerade erreichten Profilpunkt bezeichnet.

BREITE-KELLER : Der Breite-Keller dient der Berechnung von Basisbreiten von Bergen. Die Kellerelemente sind Profilpunkte. Das oberste Kellerelement wird mit $(x_B, I(x_B))$ bezeichnet, das unterste mit $(x_b, I(x_b))$. Zu Beginn ist das unterste Kellerelement ein virtueller Profilpunkt (x_o, o) (siehe unten). Für die im Keller befindlichen Punkte gilt :
$I(x_b) \leq I(x_{b+1}) \leq \ldots\ldots\ldots\ldots \leq I(x_B)$.
Dies wird durch die Operationen BVGL, BFÜLL und BLEER erreicht.

BVGL (B-Vergleich) : Durch BVGL wird geprüft ob : $I(x_B) \leq I(x_i)$ ist
Falls ja wird BFÜLL ausgeführt ;
falls nein wird BLEER ausgeführt und BVGL wiederholt.

BFÜLL (Breite-Keller-Füllen) : $(x_i, I(x_i))$ wird in den Breite-Keller geschoben.

BLEER (Breite-Keller-Leeren) : $(x_B, I(x_B))$ wird aus dem Breite-Keller geleert.

BERG-KELLER : Der Berg-Keller dient der Zwischenspeicherung von Information, um über Dominanz von Bergen entscheiden zu können. Die Kellerelemente sind Tripel (x^P, e, a). Das

oberste Tripel wird mit (x_G^P, e_G, a_G), das unterste mit (x_g^P, e_g, a_g) bezeichnet. Dabei ist x^P die X-Koordinate des Maximums, e der Tal-Index der letzten Eintragung des betreffenden Berges in der Baum-Tabelle und a die Anzahl benachbarter Berge gleicher Höhe (Sonderfall aus Abb. 13 B). Treten k solcher Berge gleicher Höhe auf, so liegen sie im Berg-Keller direkt untereinander und werden zu einer Berg-Gruppe zusammengefaßt, indem im ersten Tripel der Gruppe (Gruppenführer genannt) a=k und in den anderen Tripeln a=o gesetzt wird. Damit wird eine Berg-Gruppe wie ein einziges Kellerelement behandelt und nicht von den anderen Tripeln unterschieden. Operationen im Berg-Keller sind GFÜLL und VSCHMLZ.

GFÜLL (Berg-Keller-Füllen): ein neues Tripel (x^P, e, a) wird in den Berg-Keller geschoben.

VSCHMLZ (Vergleichen und Schmelzen) : Mit dieser Operation wird die Dominanz von Bergen entschieden und anschließend der nicht-dominante Berg aus dem Berg-Keller entfernt. Daher findet VSCHMLZ stets nach TEINTRG oder TERGNZ statt. Durch TEINTRG erfolgt ein Eintrag in die Baum-Tabelle; mit TERGNZ wird ein bestehender ergänzt (siehe unten). In VSCHMLZ werden die beiden oberen Kellerelemente des Berg-Kellers verglichen, um den dominanten Berg zu bestimmen. Es treten drei Fälle auf :

(1) Wenn $I(x_G^P) > I(x_{G-1}^P)$, dann wird $(x_{G-1}^P, e_{G-1}, a_{G-1})$ aus dem Berg-Keller entfernt.

(2) Wenn $I(x_G^P) < I(x_{G-1}^P)$, dann wird (x_G^P, e_G, a_G) aus dem Berg-Keller entfernt.

(3) Wenn $I(x_G^P) = I(x_{G-1}^P)$, dann werden beide Tripel zu einer Berg-Gruppe verschmolzen (falls eines oder beide Tripel Gruppenführer sind, werden beide Gruppen verschmolzen), wobei (x_G^P, e_G, a_G') der neue Gruppenführer wird mit : $a_G' = a_G + a_{G-1}$.

In allen drei Fällen ist e_G der Tal-Index des vorher erfolgten Eintrags in die Baum-Tabelle (durch TEINTRG oder TERGNZ).

TAL-KELLER : Der Tal-Keller dient der Erfaßung der Struktur des I-Profiles. Die Kellerelemente sind Täler. Treten benachbarte Täler gleicher Intensität auf (Sonderfall aus Abb. 13 A), dann wird die X-Koordinate des am weitesten links liegenden Tales durch eine Flagge, $\overset{*}{x}^T$, gekennzeichnet. Die restlichen Täler gleicher Intensität werden nicht in den Tal-Keller geschoben. Das oberste Kellerelement wird mit $(x_T^T, I(x_T^T))$ bezeichnet, das unterste mit $(x_t^T, I(x_t^T))$. Für die im Keller befindlichen Punkte gilt:

$I(x_t^T) \leq I(x_{t+1}^T) \leq \ldots\ldots\ldots\ldots \leq I(x_T^T)$.

Im Zusammmenhang mit dem Tal-Keller werden vier Operationen benötigt : TVGL, TFÖLL, TLEER und FLAGGE.

TVGL (T-Vergleich) : Der Profilpunkt $(x_i, I(x_i))$ wird mit dem Tal verglichen,das das oberste Kellerelement des Tal-Kellers ist. Bei diesem Vergleich muß beachtet werden, ob $(x_i, I(x_i))$ selbst ein Tal ist. In TVGL wird geprüft, ob $(x_i, I(x_i))$ auf gleicher Höhe mit $(x_T^T, I(x_T^T))$ liegt. Es treten drei Fälle auf :

(1) $I(x_T^T) < I(x_i)$:

(A) Wenn $(x_i, I(x_i))$ ein Tal ist, dann wird es in den Tal-Keller geschoben (TFÖLL).

(B) Wenn $(x_i, I(x_i))$ kein Tal ist, findet keine Veränderung im Tal - Keller statt.

(2) $I(x_T^T) = I(x_i)$:

(A) Wenn $(x_i, I(x_i))$ ein Tal ist, hat man im I-Profil einen Punkt erreicht, bei dem mehrere benachbarte Täler gleicher Intensität (Sonderfall aus Abb. 13 A) vorliegen. x_T^T wird durch eine Flagge gekennzeichnet: $\overset{*}{x}_T^T$ und eine Eintragung in der Baum-Tabelle vorgenommen (TEINTRG). Falls x_T^T bereits eine Flagge hat,bleibt diese,und die bestehende Eintragung für $(\overset{*}{x}_T^T, I(\overset{*}{x}_T^T))$ wird durch den

hinzugekommenen Berg ergänzt (TERGNZ).

(B) Wenn $(x_i, I(x_i))$ kein Tal ist, hat man einen rechten Basispunkt R (vgl. Abb. 12 A) erreicht und kann die Dominanz der beiden Berge entscheiden, deren X-Koordinaten in den oberen Tripeln des Berg-Kellers vorliegen. Hat x_T^T keine Flagge, so erfolgt ein Eintrag in die Baumtabelle (TEINTRG) mit anschließendem VSCHMLZ. Falls $\overset{*}{x}{}_T^T$ eine Flagge hat, muß ein bereits bestehender Eintrag ergänzt werden (TERGNZ).

(3) $I(x_T^T) > I(x_i)$:

(A) Dieser Fall tritt auf, weil die Profilpunkte diskret vorliegen und der Fall (2) (B) übersprungen wurde. Deshalb liegt hier stets Fall (B) vor, unabhängig davon, ob $(x_i, I(x_i))$ ein Tal ist.

(B) Es erfolgt zunächst die Behandlung des Falles wie (2)(B). Anschließend wird $(x_T^T, I(x_T^T))$ aus dem Talkeller geleert und TVGL mit $(x_i, I(x_i))$ und dem neuen oberen Kellerelement wiederholt.

TFÜLL (Tal-Keller-Füllen): es wird $(x_i, I(xi))$ in den Tal-Keller eingetragen.

TLEER (Tal-Keller-Leeren): $(x_T^T, I(x_T^T))$ wird aus dem Tal-Keller geschoben.

FLAGGE : es wird geprüft, ob x_T^T im Tal-Keller eine Flagge hat,
falls ja wird TERGNZ ausgeführt
falls nein wird TEINTRG ausgeführt

Neben den drei Kellern muß die Baum-Tabelle verwaltet werden. Dies geschieht durch die Operationen TEINTRG und TERGNZ:

TEINTRG (Tabellen-Eintrag): durch TEINTRG wird ein Knoten im Baum erzeugt. Nach Erhöhung des Tal-Index um Eins wird in die Tal-

Tabelle das Tal $(x_T^T, I(x_T^T))$ und in die Berg-Tabelle die dazugehörigen Berge (jeweils nach Erhöhung des Berg-Index) eingetragen (vgl. Abb.20).

TERGNZ (Tabellen-Ergänzung): durch TERGNZ wird der Sonderfall benachbarter Täler gleicher Intensität behandelt: eine bestehende Eintragung wird ergänzt, indem in der Berg-Tabelle der hinzugekommene Berg eingetragen wird. Diese Eintragung wird mit den bereits eingetragenen benachbarten Bergen verknüpft, indem der Nachbar-Index auf den entsprechenden Berg-Index gesetzt wird.

Sowohl bei TEINTRG als auch bei TERGNZ wird in den Eintragungen der Vorgänger der Verweis zum Tal-Index des Nachfolgers (ANC, vgl. Abb. 20) eingetragen.

Damit sind alle Operationen beschrieben, die für die Konstruktion eines R-Baumes erforderlich sind. Um alle Profilpunkte zu erfassen, müssen am linken und rechten Ende des I-Profiles noch virtuelle Profilpunkte (x_o, o) und (x_{N+1}, o) hinzugefügt werden. Abbildung 21 zeigt in einem Flußdiagramm das Zusammenspiel der oben definierten Operationen:
Der Algorithmus durchläuft das so erweiterte I-Profil von links nach rechts und befindet sich dabei in einem von zwei Zuständen: Aufwärts (AUF) oder abwärts (AB). Zu Beginn befindet er sich AUF und verbleibt darin solange $I(x_{i-1}) \leq I(x_i)$ ist. Während dieses Zustandes werden durch BFÜLL alle durchlaufenden Profilpunkte in den Breite-Keller geschoben. Wenn $I(x_{i-1}) > I(x_i)$ ist, hat der Algorithmus ein Maximum erreicht. Es wird GFÜLL und BFÜLL ausgeführt und der Zustand AB angenommen.

In AB werden die Entscheidungen getroffen, die zur Konstruktion des R-Baumes führen. Durch den Einsatz der drei Keller kann der Algorithmus gewissermaßen "nach links schauen" und so Kontext-Information ausnützen, die in den Kellern gespeichert ist. Für jeden Profilpunkt, der im Zustand AB durchlaufen wird, prüft der Algorithmus, ob es sich um ein Minimum $(I(x_i) < I(x_{i+1})?)$ handelt.

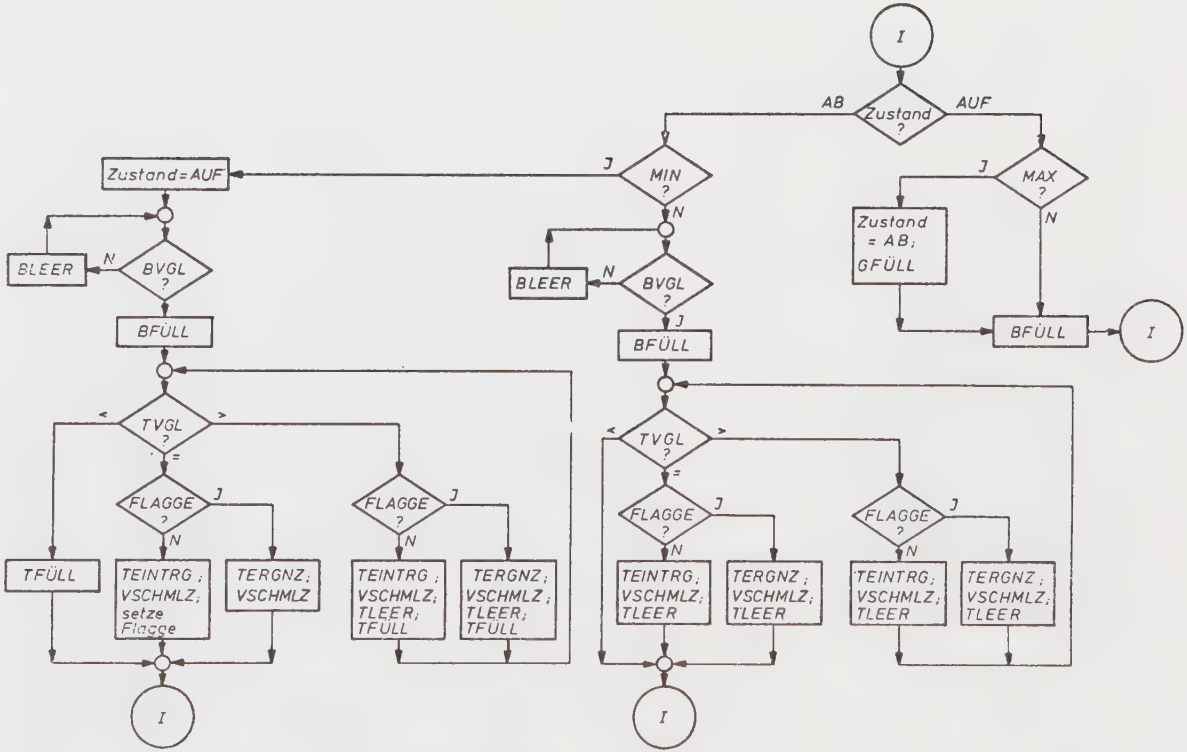


ABBILDUNG 21 : BLOCKDIAGRAMM DES ALGORITHMUS ZUR KONSTRUKTION EINES R-BAUMES

Falls $(x_i, I(x_i))$ kein Tal ist, erfolgen die Operationen BVGL zusammen mit BLEER und BFÖLL, TVGL und FLAGGE. Durch BVGL und BLEER wird der Profilpunkt gesucht, der $(x_i, I(x_i))$ auf der linken Flanke des Berges gegenüberliegt, so daß nach BFÖLL die beiden obersten Elemente des Breite-Kellers die Breite des Berges an dieser Stelle angeben. TVGL liefert einen der drei oben beschriebenen Fälle. Im Fall (1)(B) erfolgt BLEER, und BVGL wird wiederholt. Für die Fälle (2)(B) und (3)(B) erfolgt TEINTRG oder TERGNZ, je nach Ergebnis von FLAGGE. Im Fall (2)(B) erfolgen nach TEINTRG die Operationen VSCHMLZ und TLEER, anschließend wird BVGL wiederholt. Im Fall (3)(B) wird nach VSCHMLZ und TLEER die Operation TVGL wiederholt, um zu prüfen, ob mehr als ein Tal höher als $(x_i, I(x_i))$ liegt.

Falls $(x_i, I(x_i))$ ein Tal ist, wird der Zustand AUF angenommen und anschließend eine ähnliche Bearbeitung vorgenommen, wie die eben beschriebene. Die Bearbeitung unterscheidet sich in der Behandlung des Tal-Kellers: während im vorangegangenen Fall ($(x_i, I(x_i))$ kein Tal) der Tal-Keller jeweils um ein Tal geleert wird, findet hier die Operation TFÖLL statt und es werden - wenn nötig - Flaggen gesetzt.

Der bisher beschriebene Algorithmus erkennt nur Minima und Maxima. Um auch Plateaus erkennen zu können (vgl. Abschnitt 4.2), wird der Erkennungsmechanismus des Algorithmus durch einen Gradienten-Sensor erweitert. Dieser berechnet die Steigung zwischen benachbarten Profilpunkten. Bei Detektion von Minima oder Maxima wird der Gradienten-Sensor "abgeschaltet" und nach Oberschreiten des Extremwertes neu gestartet. Während der aktiven Phase werden Segmente der Steigung Null gesucht und durch einen vierten Keller, den Plateau-Keller behandelt. Kellerelemente sind die X-Koordinaten von Profilpunkten, die auf einem Plateau liegen. Bei Detektion eines Plateau-Anfangs wird der Keller solange gefüllt bis das Ende des Plateaus erreicht ist. Aus den Kellerelementen werden die beiden Endpunkte des Plateaus sowie seine Länge gewonnen und anschließend der Keller geleert. Gleichzeitig wird das Plateau als virtueller Berg in den Berg-Keller eingetragen werden. Dabei ergeben sich vier Fälle:

(1) Ein "linkes Plateau", das auf der ansteigenden Flanke eines Berges liegt. In diesem Fall wird am rechten Ende des Plateaus ein virtuelles Tal angenommen und in den Tal-Keller geschoben.

(2) Ein "flacher Berg", dessen Spitze ein Plateau ist. Dieser Berg wird wie alle anderen Berge behandelt, es wird jedoch in der Attributs-Liste vermerkt, daß der Berg oben flach ist. Die Breite des Plateaus wird in der Attributs-Liste gespeichert.

(3) Ein "rechtes Plateau", das auf der abfallenden Flanke eines Berges liegt. Hier wird ein virtuelles Tal am linken Ende des Plateaus in den Tal-Keller eingetragen.

(4) Ein "flaches Tal", d.h. ein Plateau, das auf einem Tal liegt. Ein solches Plateau wird wie ein virtueller Berg behandelt, an dessen beiden Enden je ein virtuelles Tal liegt, wobei das linke Tal eine Flagge erhält.

Mit diesen zusätzlichen Maßnahmen, die zu Beginn der Verarbeitung vorgenommen werden, läuft der verallgemeinerte Algorithmus wie der oben beschriebene ab.

Die erste Version des Baum-Algorithmus wurde auf einer PDP 15[1)] und auf einer HP 2100 A[2)] in FORTRAN IV implementiert.[3)]. Die verallgemeinerte Version wurde ebenfalls in FORTRAN IV implementiert [2), 3)].

---

1) Am Department of Electrical & Computer Engineering, Univ. of Massachusetts, Amherst, MA, USA

2) Am Image Research Lab. (Leitung: Prof. R.W. Ehrich), Dept. of Computer Science, Virginia Polytechnic Institute & State University, Blacksburg, VA, USA

3) Inzwischen wurden alle in dieser Arbeit beschriebenen Programme auf einer SIEMENS 4004 im Institut für Informationsverarbeitung in Technik und Biologie der Fraunhofer-Gesellschaft, Karlsruhe implementiert.

Eine Abschätzung des Rechenaufwandes des hier vorgestellten Konzeptes ist schwierig, da Vergleichsmöglichkeiten mit anderen Ansätzen fehlen. Das Konzept ist speicheraufwendig: für die Darstellung eines I-Profiles mit 256 Punkten in einem verallgemeinerten R-Baum werden ungefähr 3,5 K Worte zur Speicherung benötigt. Ursache des hohen Aufwandes ist eine redundante Speicherung, die die anschließende Auswertung unterstützt. So werden z.B. in der Attributs-Liste eines Berges die Koordinaten des linken und des rechten Basispunktes, des zum Berg gehörigen Tals sowie die Basisbreite gespeichert, obwohl diese Koordinaten redundant sind. Dieser erhöhte Speicheraufwand wurde zugunsten einer schnellen Verarbeitung in Kauf genommen. Eine weniger redundante Speicherung eines R-Baumes erfordert ungefähr 2 K Worte. Durch den Einsatz der Keller ist die Konstruktion der R-Bäume in einem Durchlauf durch die Profilpunkte möglich. Der Speicheraufwand für die Keller beträgt für verallgemeinerte R-Bäume ca. 2 K Worte. Rechenaufwand entsteht im wesentlichen durch Kellerbewegungen, d.h. Leeren und Füllen der Keller. Rechnet man VSCHMLZ als Leeren eines Kellers, dann finden 2(n + g + t + p) Kellerbewegungen statt, wobei n = Anzahl der Profilpunkte, g = Anzahl der Berge, t = Anzahl der Täler und p = Anzahl der Plateaus ist. Ein Beispiel für die Größenordnung der Rechenzeit, die für die Konstruktion aller R-Bäume eines Grauwertbildes mit 256 x 256 Bildpunkten und 256 Grauwerten (z.B. Abb. 11 A) benötigt wird, ist die Verarbeitungszeit auf der HP 2100 A (mit 16 K Kernspeicher), die bei 80 s liegt. Davon entfallen ca. 40 % auf Plattenzugriffszeiten, so daß die Rechenzeit für einen R-Baum in der Größenordnung von 200 ms liegt (zum Vergleich: die Berechnung des Grauwerthistogrammes für dasselbe Bild dauert auf dieser Maschine ca. 45 s.) Durch das Konzept der zeilenweisen Verarbeitung ist es möglich, stets nur diejenigen R-Bäume im Kernspeicher zu haben, die für die Auswertung relevant sind.

Ein direkter Vergleich des Aufwandes mit dem anderer Ansätze ist nicht möglich, da hierfür auch der Aufwand berücksichtigt werden muß, der bei der anschließenden Auswertung entsteht. So ist z.B. der Speicheraufwand für ein Pyramidensystem wesentlich geringer, die Aufgabe der Kantendetektion in einem Pyramidensystem jedoch aufwendiger. Beim Pyramiden-

system wird jeder Bildpunkt mehrfach von einem lokalen Operator erfaßt - und zwar in Abhängigkeit der Fenstergröße, während bei der Kantendetektion in R-Bäumen sofort und nur einmal auf die Flanken von Bergen zugegriffen wird (vgl. Abschnitt 5.2). Ein zweites Beispiel unterschiedlichen Aufwandes bei der Auswertung tritt bei Textur-Analysen auf: für die Erfassung der Größe von Flecken als Texturmaß wird ein verallgemeinerter Laplace-Operator /77/ auf das Grauwertbild angewandt, während die entsprechenden Merkmale, d.h. die Basisbreiten von Bergen, direkt den R-Bäumen entnommen werden (vgl. Abschnitt 5.4).

5. DIE AUSWERTUNG VON RELATIONAL-BÄUMEN

Mit den R-Bäumen liegt eine symbolische P-Repräsentation von Grauwertbildern vor, die Intensitäts- und Gradienten-Änderungen beschreibt. Linien, Kanten, Flecken, Regionen oder Texturen spiegeln sich in I-Profilen als charakteristische Intensitätsverläufe wieder. Die ihnen entsprechenden Teilstrukturen in den I-Profilen stellen die bedeutungsvollen Segmente dar. Diese treten auf mehreren Ebenen auf. Einfache Teilstrukturen von I-Profilen sind z.B. Segmente konstanter Intensität, solche mit konstantem Gradienten, Grauwertsprünge, usw. Typischerweise liegen solche Teilstrukturen an Konturlinien vor oder sind Teile homogener Regionen. Wir wollen allgemein solche Teilstrukturen "einfache Intensitäts-Ereignisse" (einfache I-Ereignisse) nennen. Aus der Gruppierung solcher einfacher I-Ereignisse ergeben sich zusammengesetzte Strukturen, die z.B. für Texturen typisch sind. Wir wollen solche Strukturen "komplexe I-Ereignisse" nennen. Gruppierungen können rekursiv auftreten und liefern verschiedene Beschreibungsebenen. Indem man den Ablauf dieser Gruppierungen festhält, gewinnt man Aussagen darüber, wie die verschiedenen Auswertungs-Ebenen miteinander zusammenhängen. Diese Aussagen lassen sich als Strukturen formulieren, die wir "hyperkomplexe I-Ereignisse" nennen wollen. Die Vielzahl solcher I-Ereignisse liefert ein Vokabular, aus dem sich verschiedene Auswertungs- und Beschreibungsebenen eines Bildes ergeben, die zu einer S-Repräsentation führen. Beispiele einfacher I-Ereignisse sind Kanten- und Linien-Elemente oder Regionen konstanter Intensität. Aus ihnen lassen sich zusammengesetzte Regionen als komplexe I-Ereignisse bilden. Hyperkomplexe I-Ereignisse beschreiben, wie sich die komplexen aus den einfachen I-Ereignissen zusammensetzen.

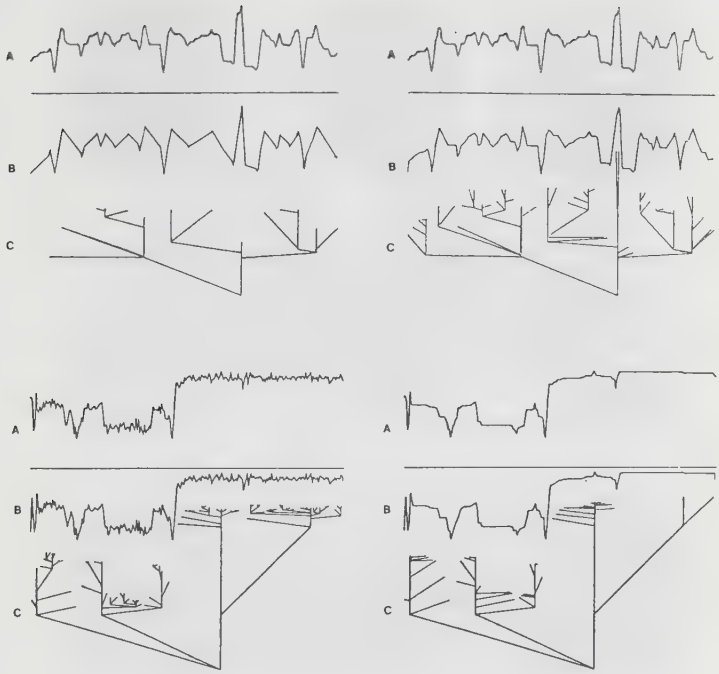
Es muß möglich sein, in den R-Bäumen Strukturen, die den I-Ereignissen entsprechen, zu finden und auszuwerten. Mit anderen Worten: die Auswertung von R-Bäumen muß eine Zerlegung in Teilstrukturen liefern, die Kanten, Linien oder Texturen entsprechen. Derartige Zerlegungen können nach verschiedenen Gesichtspunkten erfolgen. Die Zerlegungs-Kriterien müssen die Semantik der Kanten, Linien, Regionen und Texturen im Bild berück-

sichtigen, um zu sinnvollen Ergebnissen zu führen. Die Semantik dieser Elemente wird bestimmt durch die physikalische Struktur der dreidimensionalen Umwelt, die Struktur der Intensitätsverteilung im Bild und die Art, wie diese Strukturen wahrgenommen werden. Die vollständige Erfassung aller Zusammenhänge würde die Untersuchung der Beleuchtung, der Reflektionsfunktion, der physikalischen und geometrischen Bedingungen in der Umwelt und der Wahrnehmungsmechanismen einschließen müssen. Dies ist beim heutigen Stand der Technik jedoch nicht möglich. Begrenzte Aspekte dieser Zusammenhänge - z.B. zwischen Reflektionsfunktion und Geometrie der Umwelt sind für einfache Szenen untersucht /35, 36/, allgemeine Aussagen liegen jedoch nicht vor.

5.1 DIE REKONSTRUKTION VON INTENSITÄTS-PROFILEN AUS RELATIONAL-BÄUMEN

Die Rekonstruktion von I-Profilen aus den R-Bäumen stellt keine Auswertung dar, zeigt jedoch, wieviel Information durch den Übergang zu den R-Bäumen erhalten bleibt. Die Güte der Repräsentation bestimmt, wie gut die nachfolgenden Verarbeitungsschritte ausfallen. Da R-Bäume ohne Vorgabe von Fenstern oder Schwellwerten aus den I-Profilen gewonnen werden, kann ein I-Profil aus der Baumstruktur unter Berücksichtigung der Attributs-Liste zurückgewonnen werden.

Für die Rekonstruktion der I-Profile liegen zwei Programme vor: eines, das die rekonstruierten I-Profile zusammen mit der Baumstruktur grapisch auf einem Bildschirm darstellt; und ein zweites, das alle N Zeilen eines Bildes aus den R-Bäumen in $I_j(x)$, j=1,...,N überführt. Beide Programme arbeiten nach demselben Prinzip: die Koordinaten x^T , x^L , x^R, x^P der Berge und Plateaus eines R-Baumes werden von links nach rechts geordnet und die dazwischenliegenden Punkte durch lineare Interpolation ergänzt. Der Algorithmus beginnt bei der Wurzel des Baumes und läuft zum am weitesten links liegenden Blatt hoch. Knoten, die keine Blätter sind,entsprechen verschachtelten Bergen. Zur Auflösung dieser Verschachtelung wird wieder ein Keller benützt. Für jeden solchen Knoten wird der Berg-Index (vgl. Abschnitt 4.3) in den Keller geschoben, gleichzeitig zwischen $(x^L, I(x^L))$ des betreffenden Berges und dem letz-



R-BÄUME OHNE PLATEAUS R-BÄUME MIT PLATEAUS

ABBILDUNG 22 : GRAPHISCHE DARSTELLUNG DER REKONSTRUKTION VON I-PROFILEN AUS RELATIONAL-BÄUMEN

A : I-PROFIL B : REKONSTRUKTION C : R-BAUM

ten ermittelten Punkt eine Gerade gezogen. Auf diese Weise wird die linke Flanke eines Berges aufgebaut, bis ein Blatt erreicht ist. Der entsprechende Berg wird mit Geraden zwischen $(x^L, I(x^L)), (x^P, I(x^P))$ und $(x^R, I(x^R))$ rekonstruiert. Der Algorithmus geht zum nächsten Knoten im Baum zurück, von dem aus das nächste Blatt erreicht werden kann. Während der Rücksprünge wird der Keller geleert und die rechte Flanke des Berges aus den $(x^R, I(x^R))$ aufgebaut, die in der Attributs-Liste mit dem gerade geleerten Berg-Index gespeichert sind. Wird anstelle der graphischen Darstellung die Rekonstruktion von Bildzeilen gewählt, so wird den Bildpunkten x_i der gespeicherte bzw. interpolierte Wert $I(x_i)$ zuge-

ABBILDUNG 23 : DIE REKONSTRUKTION VON GRAUWERTBILDERN AUS RELATIONAL-BÄUMEN

A - ORIGINAL

B - REKONSTRUKTION OHNE PLATEAUS

C - REKONSTRUKTION MIT PLATEAUS

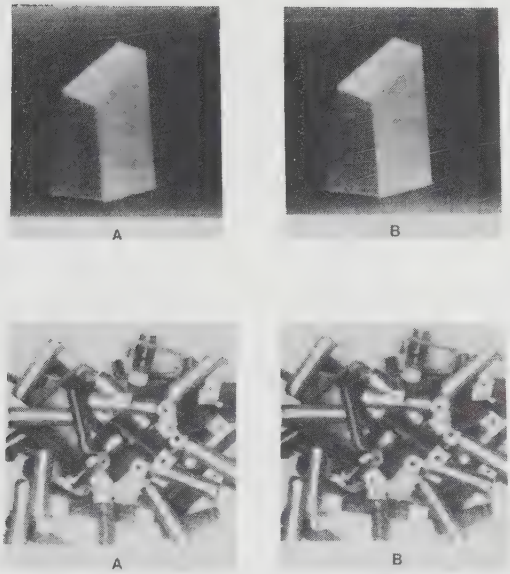


ABBILDUNG 24 : DIE REKONSTRUKTION VON GRAUWERTBILDERN AUS VERALLGEMEINERTEN R-BÄUMEN

A - ORIGINAL B - REKONSTRUKTION MIT PLATEAUS

wiesen. In beiden Fällen liegt nach einmaligem Durchlauf des Baumes die Rekonstruktion vor.

Abbildung 22 zeigt die graphische Darstellung der Rekonstruktion an einigen Beispielen. Werden R-Bäume ohne Plateaus verwendet, dann führt die Rekonstruktion des I-Profiles zu Verzerrungen, da Plateaus teilweise durch steile Flanken ersetzt werden. Durch die Erfassung von Plateaus in den verallgemeinerten R-Bäumen wird eine wesentliche Verbesserung erreicht,und die Rekonstruktionen kommen sehr nahe an die ursprünglichen I-Profile heran. (Bei den in Abb. 22 gezeigten I-Profilen handelt es sich um geglättete Daten).

Die Abbildungen 23 und 24 zeigen Beispiele rekonstruierter Grauwertbilder. Die Verzerrungen, die bei R-Bäumen ohne Plateaus entstehen, treten in den Grauwertbildern als Streifen auf, die besonders in den homogenen Regionen der Bilder zu beobachten sind. Die Rekonstruktion der Grauwertbilder aus verallgemeinerten R-Bäumen führt hingegen zu brauchbaren Bildern.

5.2 KANTENDETEKTION

In Grauwertbildern sind Kanten - und Linien-Elemente einfache I-Ereignisse. Wie sehen die Strukturen der I-Profile aus, die diesen einfachen I-Ereignissen entsprechen? Vielen Kantendetektoren der Literatur liegen Kanten-Modelle zugrunde, die eine Stufe im I-Profil darstellen, während Linien-Elemente durch Spitzen oder Dächer modelliert werden (vgl. Abb. 2). In natürlichen Szenen treten diese einfachen I-Ereignisse nicht in reiner Form auf, sondern sind verzerrt und häufig durch störende Strukturen überlagert.

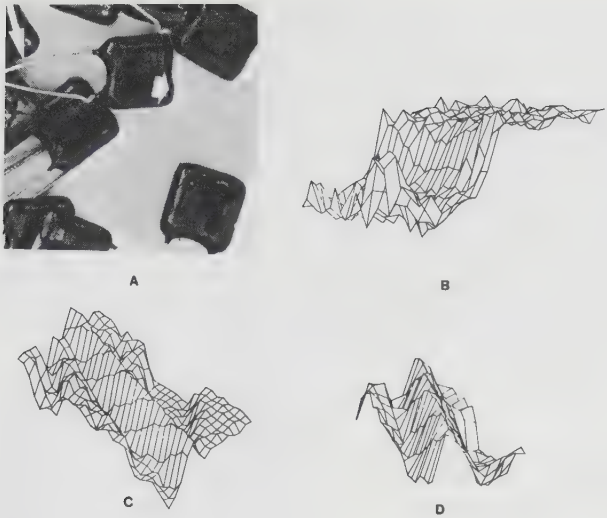


ABBILDUNG 25 : DIE STRUKTUR VON KANTEN UND LINIEN IM GRAUWERT-GEBIRGE

A - GRAUWERTBILD
B - KANTE (DICKER PFEIL) ; UM 30^0 NACH OBEN GEDREHT
C - LINIE (DÜNNER PFEIL) ; UM 45^0 NACH UNTEN GEDREHT
D - LINIE (DÜNNER PFEIL) ; UM 70^0 NACH UNTEN GEDREHT

Abbildung 25 zeigt ein Grauwertbild (Abb. 25 A), in dem scharfe Kanten und Linien vorkommen. Die Struktur der Kanten und Linien im Grauwertgebirge wird an je einer Kante und Linie durch eine dreidimensionale Darstellung der Grauwertverteilung verdeutlicht. Die Beispiele sind in Abb. 25 A durch Pfeile gekennzeichnet. In der Struktur der Kante sind Absätze und Vorsprünge zu beobachten, die eine exakte Plazierung eines Kantenelementes erschweren (Abb. 25 B). Bei der Linie wird ein anderes Problem deutlich: durch die wechselnde Umgebung, in die die Linie eingebettet ist, treten an den Seiten der Linie unterschiedliche Kontrastverhältnisse auf, die eine exakte Erfassung der Linie schwierig machen (Abb. 25 C, D). Verzerrungen und der Einfluß von Kontext führen beim Einsatz lokaler Operatoren, wie die in Abschnitt 3.1 und 3.2 beschriebenen, zu verfälschten Werten (Fensterproblem) und zu mehrdeutigen Plazierungen.

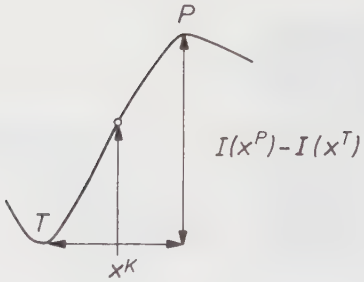


ABBILDUNG 26 : DIE BESTIMMUNG EINES KANTENELEMENTES AN DER FLANKE EINES BERGES

Kanten (und Linien) können aus R-Bäumen extrahiert werden, da in den R-Bäumen die Struktur der I-Profile gespeichert ist. Gleichzeitig kann Kontext-Information über benachbarte I-Ereignisse ausgewertet werden. Ein einfaches Beispiel der Kantendetektion in R-Bäumen ist in Abbildung 26 dargestellt. Kanten treten in R-Bäumen an den Flanken der Berge/Plateaus auf. Zu ihrer Ermittlung werden alle Berge von links nach rechts geordnet (wie bei der Rekonstruktion, vgl. Abschnitt 5.1) und die Flanken zwischen Minima und Maxima ausgewertet. Sei $T=(x^T,I(x^T))$ der linke und $P=(x^P,I(x^P))$ der rechte Endpunkt einer Flanke. Dann wird als Plazierung x^K des Kantenelements die Mitte zwischen x^T und x^P gewählt:

$$x^K = (x^T + x^P)/2.$$

Als Stärke $ST(x^K)$ des Kantenelementes bezeichnet man die relative Höhe:

$$ST(x^K) = I(x^P) - I(x^T).$$

Da auf diese Weise einfache Teilstrukturen von I-Profilen ausgewertet werden, wird dieser Detektor als "struktureller Kantendetektor" bezeichnet. Abbildung 27 C zeigt das Ergebnis dieses Kantendetektors, wobei R-Bäume sowohl in Zeilen- als auch Spalten-Richtung berechnet und ausgewertet wurden. Zum Vergleich wird das Ergebnis eines herkömmlichen Kantendetektors danebengestellt (Abb. 27 B). Es handelt sich um den "Sobel"-Operator $S(x_i,y_j)$, der in Abschnitt 3.2 beschrieben wurde. Der Vergleich der Ergebnisse zeigt folgendes: der "Sobel"-Operator findet nicht alle Konturen - z.B. in den Wolken oder im Gras; die dicken

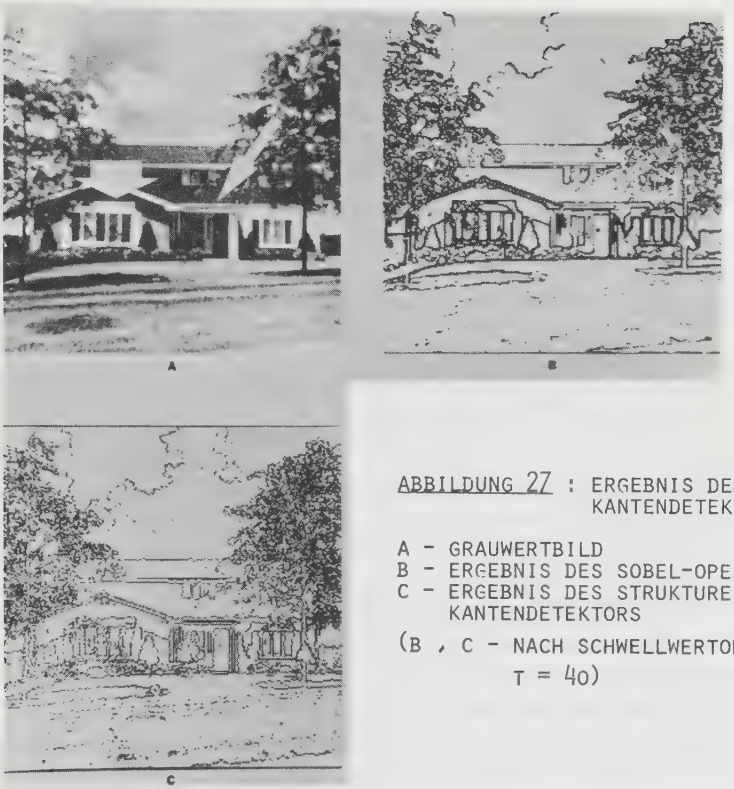


ABBILDUNG 27 : ERGEBNIS DER KANTENDETEKTION

A - GRAUWERTBILD
B - ERGEBNIS DES SOBEL-OPERATORS
C - ERGEBNIS DES STRUKTURELLEN KANTENDETEKTORS

(B , C - NACH SCHWELLWERTOPERATION T = 40)

Linien der Konturen in Abb. 27 B sind eine Folge mehrfacher Plazierungen von Kantenelementen für eine einzige Kante - d.h. es werden mehrere benachbarte Bildpunkte als Kantenelement ermittelt, obwohl nur eine Kante vorliegt; Linien müssen durch Kantendetektoren in zwei parallele Kanten aufgelöst werden - der "Sobel"-Operator ist dazu nicht in der Lage (ein Beispiel ist die schwarze Linie in Abb. 27 A, die durch einen dünnen Pfeil gekennzeichnet ist). Beim strukturellen Kantendetektor treten diese Probleme nicht auf. Die Ursache der Probleme, die beim "Sobel"-Operator auftreten, wird aus Abbildung 28 ersichtlich. Diese Abbildung zeigt die dreidimensionale Struktur zweier Bildausschnitte aus Abb. 27, die dort durch Pfeile gekennzeichnet sind. Abb. 28 A gibt die Struktur einer Linie wieder (dünner Pfeil in Abb. 27 A), die durch den "Sobel"-

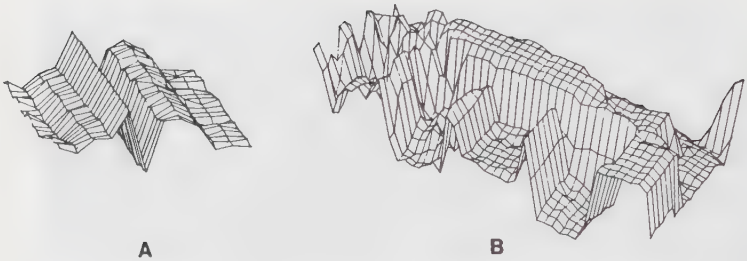


ABBILDUNG 28 : DREIDIMENSIONALE DARSTELLUNG DES GRAUWERT-GEBIRGES IN BILDAUSSCHNITTEN AUS ABB. 27

A - BILDAUSSCHNITT DÜNNER PFEIL
(DARSTELLUNG UM 80^0 NACH UNTEN GEDREHT)

B - BILDAUSSCHNITT DICKER PFEIL
(DARSTELLUNG UM 30^0 NACH UNTEN GEDREHT)

Operator nicht aufgelöst wird. Durch das 3 x 3 - Fenster, auf dem der "Sobel"-Operator definiert ist, wird von den beiden Kanten der Linie nur eine detektiert. Der strukturelle Kantendetektor findet hingegen beide Kanten. Bei dem zweiten Bildausschnitt (dicker Pfeil in Abb. 27 A) liegt das entgegengesetzte Problem vor: es werden an einer Kante mehrere benachbarte Bildpunkte als Kantenelemente ermittelt, so daß in Abb. 27 B die entsprechenden Konturen verschmiert sind. Abb. 28 B zeigt, daß die Kanten sich über mehrere Bildpunkte erstrecken, die alle in einem Gebiet mit hohem Gradienten liegen. Deshalb liefert der "Sobel"-Operator für jeden dieser Punkte ein Kantenelement.

Während beim "Sobel"-Operator jeder Bildpunkt achtmal in die Berechnung des Gradienten eingeht, werden beim strukturellen Kantendetektor nicht einzelne Bildpunkte erfaßt, sondern nur Täler und Berge - und zwar je einmal. Das Durchlaufen der R-Bäume entspricht einer Folge von Zuweisungen von Adressen, so daß nur bei der Bestimmung von x^K und $ST(x^K)$ gerechnet werden muß. Geht man davon aus, daß ein Grauwertbild bereits in einer P-Repräsentation in Form von R-Bäumen vorliegt, dann ist der Rechenaufwand des strukturellen Kantendetektors wesentlich geringer als der Aufwand des herkömmlichen Kantendetektors (ca. 4:1). Ein weiterer Vorteil der Kantendetektion in R-Bäumen wird beim Vergleich mit dem Ansatz der "Erst-Skizze" ('Primal Sketch'/44/) deutlich: das symbolische Vokabular der "Erst-Skizze" beschreibt die Struktur von

Grauwertänderungen und wird durch lokale Operatoren erzeugt. Um Aussagen über die Struktur von Kanten- und Linien-Elementen zu gewinnen, werden an jedem Bildpunkt mehrere Operatoren eingesetzt, aus deren Ergebnissen das "beste" ausgewählt wird. Das Verfahren liefert trotz hohem Rechenaufwand nur Schätzungen über die Struktur der gesuchten Elemente. Im Gegensatz dazu kann die in den R-Bäumen vorhandene Strukturinformation ohne großen Aufwand für die Erzeugung eines symbolischen Vokabulars verwendet werden. Durch ein solches Vokabular können einfache I-Ereignisse beschrieben werden. Dabei liefern die R-Bäume auch Kontext-Information, die die weitere Verarbeitung unterstützen kann. Abbildung 29 enthält ein Beispiel, wie ein Element eines symbolischen Vokabulars einfacher I-Ereignisse aussehen könnte. Es ist klar, daß die Implementierung eines solchen Vokabulars sehr hohen Speicher-Aufwand erfordert. Dafür ist zu erwarten, daß die nachfolgende Verarbeitung - z.B. die Linienverfolgung - auf diesem Vokabular wesentlich bessere Ergebnisse liefert.

Linienelemente entsprechen Bergen geringer Basisbreite bzw. dicht benachbarten Kantenelementen (vgl. Abb. 28 A). Im ersten Fall liegen helle Linienelemente auf dunklem Grund vor. Diese können durch Auswahl der Berge aus den Blättern des R-Baumes ermittelt werden. Im zweiten Fall liegen dunkle Linienelemente vor, die aus parallelen Kantenelementen zusammengesetzt werden. Auch Linienelemente können durch ein Vokabular beschrieben werden (vgl. Abb. 51).

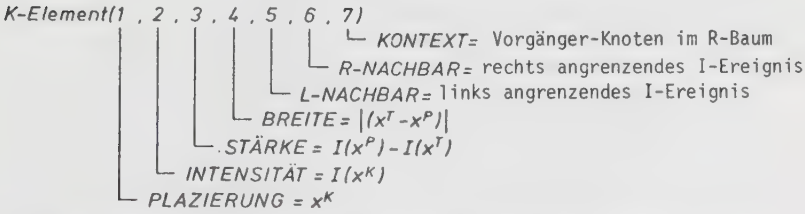


ABBILDUNG 29 : ELEMENT EINES SYMBOLISCHEN VOKABULARS ZUR BESCHREIBUNG EINFACHER I-EREIGNISSE

HIER : EIN KANTEN-ELEMENT (K-ELEMENT)

5.3 GRUPPIERUNGS-MECHANISMEN

Leistungsfähige niedere Prozesse erfordern auch höhere Beschreibungsebenen als die der einfachen I-Ereignisse. Wie gelangt man zu diesen Ebenen? Der Übergang zur nächsten höheren Ebene kann z.B. durch Gruppierung der einfachen Elemente einer Ebene zu komplexeren Elementen bewirkt werden. Aus der Vielzahl der möglichen Gruppierungs-Mechanismen sollen hier zwei erörtert werden, die in R-Bäumen leicht zu realisieren sind.

5.3.1 GRUPPIERUNGEN

Die Grundelemente von R-Bäumen sind Berge. Eine Möglichkeit der Gruppierung von Bergen besteht darin, benachbarte Berge zu Gruppen zusammenzufassen, die auf der höheren Ebene wieder als Berge auftreten. Dies trifft bei den im folgenden beschriebenen Gruppierungen zu, wobei unterschiedliche Kriterien festlegen, welche Berge zu einer Gruppe gehören.

DIE SPLITTER-GRUPPIERUNG

Aufgabe der Splitter-Gruppierung ist es, benachbarte Berge zu verschmelzen, die zu einem "Splitter" gehören. Ein Splitter ist eine Gruppe benachbarter Berge, deren begrenzende Täler zunächst aufsteigen und dann streng abfallen, so daß eine Verbindungslinie zwischen allen Tälern eines Splitters wieder einen Berg ergibt (Abbildung 30).

Definition: Sei x^{T_i} die X-Koordinate des i-ten Tals und ferner $B_k = (P_i, \ldots, P_k, \ldots, P_l)$ eine Berg-Gruppe, so daß:

(1) $I(x^{T_i-2}) > I(x^{T_i-1})$

(2) $I(x^{T_{i-1}}) \leq I(x^{T_i}) \leq \ldots\ldots \leq I(x^{T_k})$

(3) $I(x^{T_k}) > I(x^{T_{k+1}}) > \ldots\ldots > I(x^{T_l})$

(4) $I(x^{T_l}) < I(x^{T_{l+1}})$

Ein Splitter ist ein Teilbaum eines Relational-Baumes und besteht aus allen Blättern, die den Bergen in B_k entsprechen sowie allen Knoten und Kanten, die notwendig sind, um diese Blätter miteinander zu verbinden. Die Wurzel des Splitters ist derjenige Knoten des Splitters, der der Wurzel des R-Baumes am nächsten liegt.

Splitter berücksichtigen nur die Folge der lokalen Minima. Da die Struktur der R-Bäume durch Folgen von Minima bestimmt wird, ist es möglich, Splitter allein durch Analyse der Baumstruktur zu ermitteln. Der Algorithmus zur Bestimmung der Splitter beruht auf dem Durchlaufen von Knoten im R-Baum nach folgenden Regeln (siehe Abbildung 30 A):

(1) Beginne beim am weitesten links liegenden Blatt und gehe zu seinem Vorgänger. Markiere durchlaufene Knoten mit einer Splitter-Nummer i.

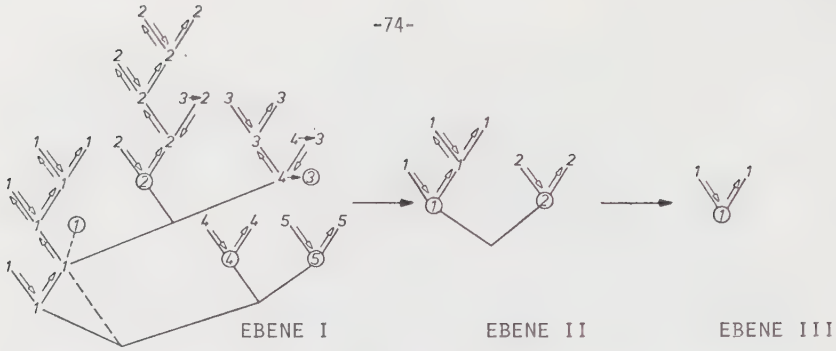
(2) Gehe zum am weitesten links liegenden Nachfolger, der noch nicht markiert ist. Wiederhole (2), bis ein Blatt erreicht ist.

(3) Nach Erreichen eines Blattes gehe zu dessen Vorgänger zurück und wiederhole (2).

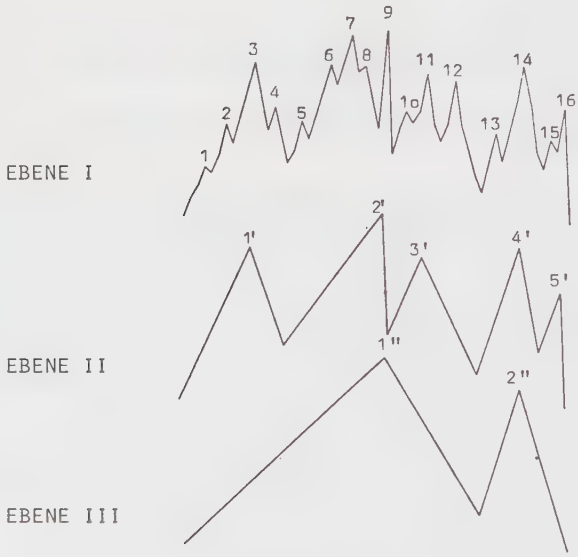
(4) Das Ende eines Splitters ist erreicht, wenn nach einem Rückschritt der Länge 1 (Regel (3)) kein unmarkiertes Blatt mehr gefunden wird.

(5) Stößt man beim Markieren mit Marke j auf einen Knoten mit Marke i, so hat man einen unvollständigen Splitter erreicht und muß alle mit j markierten Knoten in i ummarkieren.

Sind alle Splitter eines R-Baumes ermittelt, bleibt die Aufgabe, das Ergebnis der Gruppierung auf der nächsten Ebene darzustellen. Eine Möglichkeit bietet eine Darstellung, bei der die Berg-Gruppe eines Splitters durch einen einzigen Berg ersetzt wird (Abbildung 30 B). In den R-Bäumen entspricht dies dem Entfernen aller Knoten eines Splitters mit Ausnahme der Splitter-Wurzel. Einen Sonderfall bilden Knoten eines Splitters, deren Nachfolger nicht alle zu dem betreffenden Splitter ge-



A - REKURSIVE ANWENDUNG DER SPLITTER-GRUPPIERUNG : NACH ERMITTLUNG DER SPLITTER EINES R-BAUMES WERDEN VON JEDEM SPLITTER ALLE KNOTEN - MIT AUSNAHME DER SPLITTERWURZEL (DURCH ○ GEKENNZEICHNET) - ENTFERNT. AUF DEN ENTSTANDENEN R-BAUM WIRD WIEDER DIE GRUPPIERUNG ANGEWANDT. (DIE KNOTEN DER R-BÄUME SIND HIER MIT DER BETREFFENDEN SPLITTER-NUMMER MARKIERT.)



B - DIE WIRKUNG DER REKURSIVEN ANWENDUNG DER SPLITTER-GRUPPIERUNG AUF I-PROFILE

<u>ABBILDUNG 30</u> : ÜBERGÄNGE ZWISCHEN BESCHREIBUNGS-EBENEN, DIE DURCH DIE SPLITTER-GRUPPIERUNG ERZEUGT WERDEN

A - IN DEN R-BÄUMEN B - IN DEN I-PROFILEN

hören, d.h. Knoten, von denen aus andere Splitter erreicht werden können. Solche Knoten dürfen nicht entfernt werden, da sie einem Tal entsprechen, das zwischen zwei Splittern liegt (vgl. Abb. 30 A, Ebene I, Splitter 1).

Abbildung 31 zeigt Grauwertbilder, die durch die rekursive Anwendung der Splitter-Gruppierung mit der oben beschriebenen Art der Darstellung entstanden sind. Da bei der Splitter-Gruppierung lediglich die Relation von Tälern eine Rolle spielt, wird Information über Berge nicht erfaßt. Es kommt daher darauf an, ob die bedeutungsvollen Regionen im Bild als helle Figuren auf dunklem Hintergrund vorliegen oder umgekehrt. Im ersten Fall treten die hellen Regionen als Berge auf und werden entsprechend gruppiert. Im zweiten Fall muß das I-Profil invertiert werden, damit die dunklen Regionen als Berge erscheinen. Dies wird aus Abb. 31 deutlich. In Abb. 31 A werden die hellen Regionen gruppiert. Dabei wachsen die hellen Regionen nach mehrfacher Anwendung der Splitter-Gruppierung an einigen Stellen an, während die dunklen Regionen langsam grau werden. In Abb. 31 B ist die Verarbeitung des invertierten Bildes gezeigt (d.h. Grauwert 255 = Schwarz, Grauwert 0 = Weiß). Hier treten die dunklen Kondensatoren als Berge auf. In diesem Fall tritt die beabsichtigte Wirkung der Splitter-Gruppierung ein: die Regionen füllen sich langsam auf und verschmelzen zu homogenen Regionen. Auf diese Weise bildet die Splitter-Gruppierung ein Verfahren zur Flächenbildung.

Durch eine Darstellung, bei der alle Berge eines Splitters durch einen neuen Berg ersetzt werden, entstehen häufig Verzerrungen, die sich als Streifen im Grauwertbild bemerkbar machen. Eine konservative Art der Darstellung bietet hier Abhilfe: man bildet eine Hüllkurve eines Splitters, indem man die Maxima aller Berge innerhalb eines Splitters durch Geraden verbindet. Dabei können innerhalb eines Splitters neue Täler entstehen, die bei rekursiver Anwendung dieser modifizierten Splitter-Gruppierung zu anderen Ergebnissen als bei der ursprünglichen Gruppierung führen.

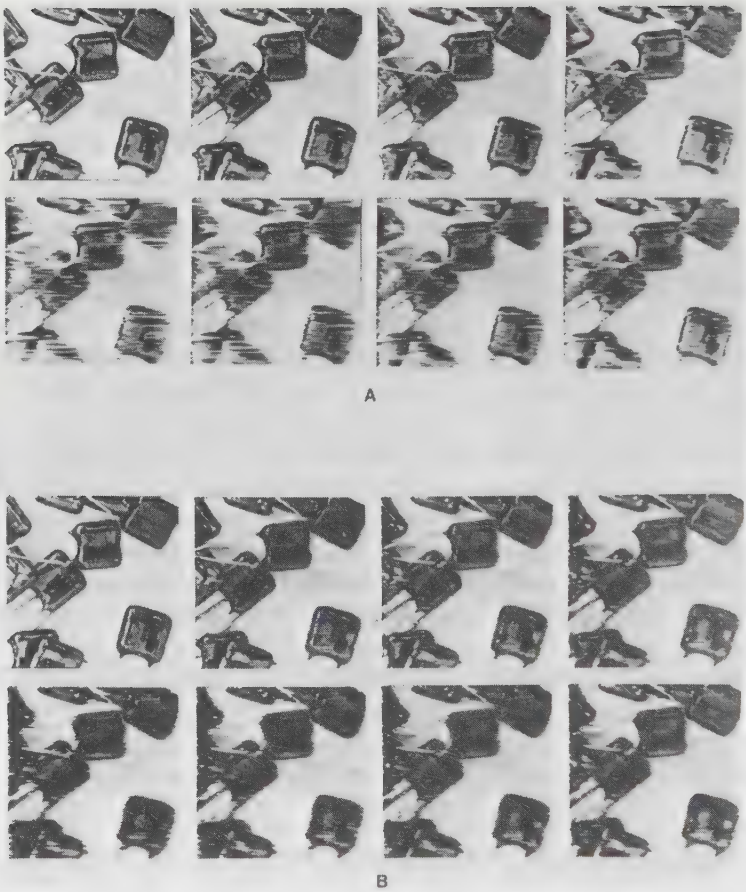


ABBILDUNG 31 : REKURSIVE ANWENDUNG DER SPLITTER-GRUPPIERUNG IN ZEILENRICHTUNG

A - HELLE REGIONEN ALS FIGUR B - DUNKLE REGIONEN ALS FIGUR

(DARSTELLUNG IM UHRZEIGERSINN, LINKS OBEN BEGINNEND)

DIE STAMM-GRUPPIERUNG

Auch die "Stamm-Gruppierung" dient dem Verschmelzen benachbarter Berge. Die Stamm-Gruppierung bildet eine feinere Zerlegung der R-Bäume als die Splitter-Gruppierung, da hier auch Information über die Dominanz von Bergen berücksichtigt wird. Ein "Stamm" ist ein dominanter Berg, der einige Nebenberge an den Flanken hat, wobei die Täler an den Flanken nach rechts und links streng fallende Intensitäten haben (siehe Abbildung 32).

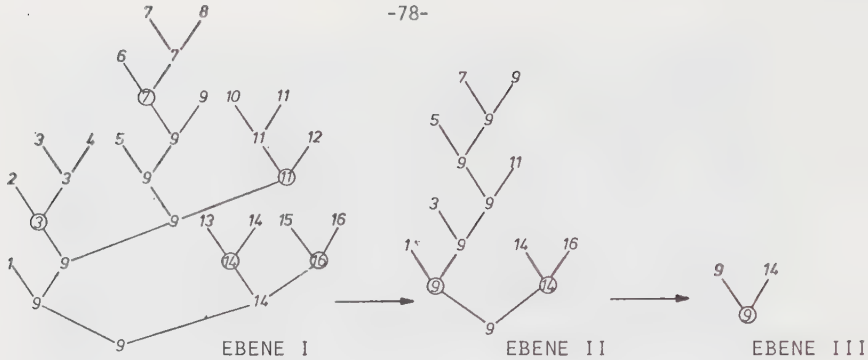
DEFINITION: Der Weg eines Stammes in einem R-Baum besteht aus Knoten mit identischen Marken. Ein Stamm besteht aus den Knoten, die auf dem Weg des Stammes liegen und allen Nachfolgern einschließlich Blättern dieser Knoten. Aus Gründen der Eindeutigkeit darf kein Stamm Knoten eines anderen Stammes enthalten.

Die Stamm-Gruppierung wird mit Hilfe der Splitter-Gruppierung ausgeführt. Neben der Baumstruktur werden auch die Marken der Knoten des R-Baumes berücksichtigt:

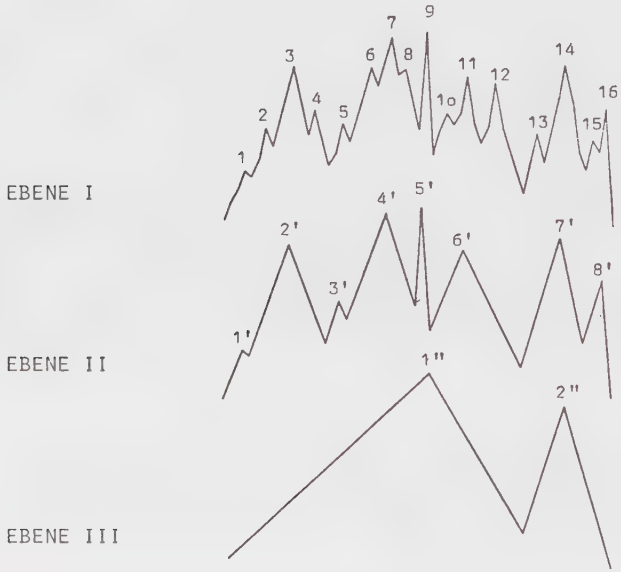
(1) Zuerst werden die Splitter des R-Baumes ermittelt.

(2) Innerhalb jedes Splitters werden die Wege von Stämmen bestimmt (mit Länge > 1). Aus diesen Wegen wird der kürzeste Weg ausgewählt und der entsprechende Stamm gebildet. Als Wurzel des Stammes wird derjenige Knoten des Stamm-Weges gewählt, der der Splitter-Wurzel am nächsten liegt (in Abb. 32 durch Kreise gekennzeichnet).

Wie bei der Splitter-Gruppierung bestehen zwei Möglichkeiten, die Ergebnisse darzustellen: entweder werden alle Berge eines Stammes durch einen neuen Berg ersetzt (vgl. Abb. 32 B) oder alle Maxima der Berge eines Stammes durch Geraden verbunden.



A - REKURSIVE ANWENDUNG DER STAMM-GRUPPIERUNG : NACH ERMITTLUNG DER STÄMME EINES R-BAUMES WERDEN VON JEDEM STAMM ALLE KNOTEN - MIT AUSNAHME DER STAMMWURZEL (DURCH ○ GEKENNZEICHNET) - ENTFERNT. AUF DEN ENTSTANDENEN R-BAUM WIRD WIEDER DIE GRUPPIERUNG ANGEWANDT. (DIE KNOTEN DER R-BÄUME SIND IN DER ÜBLICHEN WEISE -D.H. DURCH BERGE - MARKIERT.)



B - DIE WIRKUNG DER REKURSIVEN ANWENDUNG DER STAMM-GRUPPIERUNG AUF I-PROFILE

<u>ABBILDUNG 32</u> : ÜBERGÄNGE ZWISCHEN BESCHREIBUNGS-EBENEN, DIE DURCH DIE STAMM-GRUPPIERUNG ERZEUGT WERDEN

A - IN DEN R-BÄUMEN B - IN DEN I-PROFILEN

Abbildung 33 zeigt an einem einfachen Beispiel die Unterschiede der verschiedenen Darstellungen, und zwar für jede der beiden Gruppierungen.

DIE WIRKUNG VON GRUPPIERUNGEN

Die beiden hier besprochenen Gruppierungen bewirken, daß benachbarte Berge so angeordnet werden, daß wieder Berge entstehen. Auf diese Weise wird die Verschachtelung von Bergen aufgelöst. Abbildung 34 verdeutlicht diese Wirkung. In Abbildung 34 A liegt ein Grauwertbild einer Textur

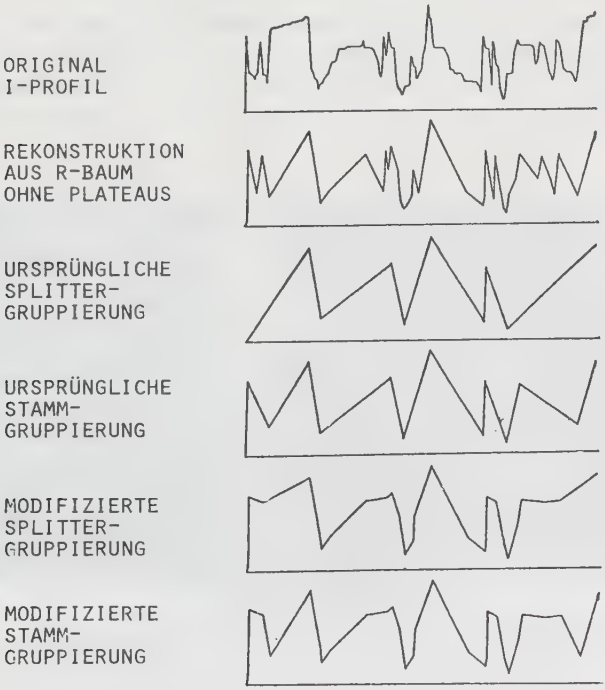
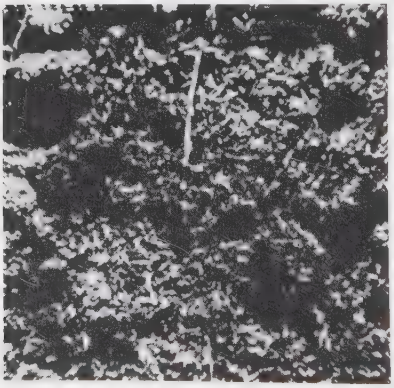


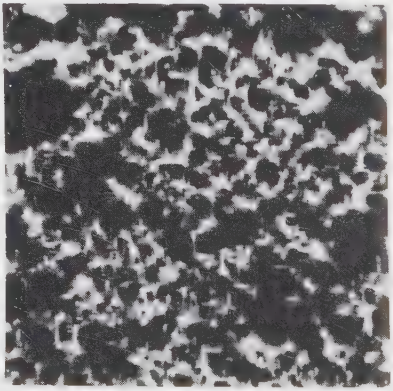
ABBILDUNG 33 : EIN VERGLEICH DER DARSTELLUNGEN DER ERGEBNISSE NACH SPLITTER- BZW. STAMM-GRUPPIERUNGEN

(Moos) vor. Abbildung 34 B zeigt das Ergebnis einer rekursiven Anwendung der Splitter-Gruppierung (7 Iterationen). Für jede Iteration wurden die R-Bäume in Zeilen- und Spaltenrichtung konstruiert und auf diese die Splitter-Gruppierung angewandt. Die Ergebnisse der zeilen- und spaltenweisen Verarbeitung wurden durch Berechnung des Mittelwertes beider Ergebnisse zusammengelegt und als neues Grauwertbild ausgegeben. Aufgrund der Gruppierung bilden die hellen Regionen Häufungsgebiete um lokale Maxima der Grauwertverteilung. Die entstehenden Regionen sind relativ homogen, obwohl das Grauwertbild in Abbildung 34 A stark texturiert ist. Die derart gebildeten homogenen Regionen können als Keimzellen für eine weitere Verarbeitung herangezogen werden.

Viele Verfahren der Flächenbildung beruhen darauf, benachbarte Regionen zu verschmelzen, deren Merkmale (z.B. mittlerer Grauwert) sich in ihren Werten höchstens um einen Betrag ϑ unterscheiden. Bei der Splitter- und Stamm-Gruppierung hingegen werden Berge nicht aufgrund ähnlicher



A - GRAUWERTBILD EINER TEXTUR



B - GRUPPIERUNG VON TEXTURELEMENTEN ZU HOMOGENEN REGIONEN

ABBILDUNG 34 : DIE WIRKUNG DER SPLITTER-GRUPPIERUNG NACH MEHRFACHER ANWENDUNG IN ZEILEN- UND IN SPALTEN-RICHTUNG

Werte ihrer Attribute miteinander verknüpft. Statt dessen wird die Entscheidung darüber, zu welcher Berg-Gruppe ein Berg gehört, allein aufgrund der Strukturen getroffen, in die der betreffende Berg eingebettet ist. Verfahren, bei denen Regionen mit ähnlichen mittleren Grauwerten verschmolzen werden, gelingt es nicht, die in Abb. 34 B gezeigten Regionen zu bilden (vgl. Abb. 38).

Gruppierungen, bei denen Ähnlichkeiten als Kriterium dienen, sind ebenso notwendig, wie die hier erörterten Gruppierungen. Bei derartigen Gruppierungen können beispielsweise benachbarte Berge mit ähnlichen Attributen miteinander verschmolzen werden. Auch hier können Strukturen, in die die Berge eingebettet sind, zur Entscheidung herangezogen werden. Solche Gruppierungen sind in den R-Bäumen grundsätzlich möglich; es wurde jedoch nicht versucht, konkrete Algorithmen zu entwickeln.

Werden die Splitter- oder Stamm-Gruppierung rekursiv angewandt, so führt dies dazu, daß Knoten aus den R-Bäumen entfernt werden. Dies entspricht der Auswahl von Knoten, die in der Nähe der Wurzel liegen und große "globale" Berge repräsentieren. Die Auswertung von Knoten nahe der Wurzel erlaubt eine Auswertung der Grobstrukturen des Bildes und stellt

ABBILDUNG 35 : GROBSTRUKTUREN DURCH AUSWERTUNG VON KNOTEN NAHE DEN WURZELN VON R-BÄUMEN

damit eine Alternative zum Konzept der "Pläne" in Pyramidensystemen dar (vgl. Abschnitt 3.2). Abbildung 35 zeigt, wie solche Grobstrukturen aussehen, wenn Knoten in der Nähe der Wurzel ausgewertet werden (vgl. Abbildung 4). Die Abbildung zeigt eine mehrfache Anwendung der Stamm-Gruppierung, wobei für die Darstellung des Ergebnisses jeweils der mittlere Grauwert innerhalb eines Stammes gewählt wurde. Für die Auswahl geeigneter Knoten nahe der Wurzel lassen sich eine Reihe anderer Auswahlverfahren denken, bei denen sowohl die Attribute der Berge als auch die Baumstruktur für Entscheidungen ausgenützt werden. Darauf wird jedoch nicht näher eingegangen.

Die oben eingeführten Techniken der Gruppierung sind sicher nicht die endgültige Lösung für Methoden der Flächenbildung. Die Experimente demonstrieren jedoch, daß die Gruppierungen grundsätzlich in der Lage sind, Regionen zu erfassen (vgl. Abbildung 31 und 34), auch wenn Texturen vorliegen. Eine Schwäche dieser Verfahren beruht darin, daß nur Informationen in Zeilenrichtung ausgewertet werden. Die Ergänzung dieser Verfahren durch eine Verarbeitung quer zu den Zeilen wird in Abschnitt 5.5 erörtert.

5.3.2 TIEFENSTRUKTUREN

Durch Anwendung der Splitter- oder Stamm-Gruppierung wird eine Beschreibungsebene von I-Profilen in eine andere Ebene überführt, auf der nicht mehr einzelne Berge, sondern nur noch Berg-Gruppen unterschieden werden. Bei rekursiver Anwendung der Gruppierungen wiederholt sich dieser Vorgang. Der Zusammenhang zwischen den verschiedenen Beschreibungsebenen kann erfaßt werden, indem man den Ablauf der Gruppierungen festhält. Abbildung 36 erklärt dieses Konzept an einem Beispiel (vgl. auch Abb.30). Durch Anwendung einer Splitter-Gruppierung werden die Berge des I-Profils aus Abb. 30 B zu den Berg-Gruppen:
$(P_1P_2P_3P_4)$, $(P_5P_6P_7P_8P_9)$, $(P_{10}P_{11}P_{12})$, $(P_{13}P_{14})$ und $(P_{15}P_{16})$ zusammengefaßt, so daß sie auf der nächsten Ebene als P'_1 P'_2 P'_3 P'_4 P'_5 auftreten.

ABBILDUNG 36 : BEISPIEL EINES STRUKTUR-BAUMES
A : INTENSITÄTS-PROFIL
B : STRUKTUR-BAUM FÜR EINE DREIMALIGE SPLITTER-GRUPPIERUNG

Bei Wiederholung der Splitter-Gruppierung werden die Berg-Gruppen (P_1' P_2' P_3') und (P_4' P_5') gebildet und man erhält auf der nächsten Ebene P_1'' P_2''. Durch diese Zusammenhänge entsprechen also dem Berg P_1'' auf der Ebene III die Berge P_1 ..... P_{12} auf der Ebene I.

Diese Zusammenhänge können durch einen Baum dargestellt werden, der "Struktur-Baum" (S-Baum) genannt werden soll (vgl. Abb. 36 B).

Die Struktur des S-Baumes enthält somit Information über die Übergänge zwischen Beschreibungsebenen und stellt eine "Tiefenstruktur" dar, die sich erst aus der Betrachtung mehrerer Ebenen ergibt. Da diese Strukturen Aussagen über Zusammenhänge zwischen komplexen I-Ereignissen

machen, wollen wir sie "hyperkomplexe I-Ereignisse" nennen. Sie beschreiben, wie komplexe I-Ereignisse zustandekommen.

Bei Anwendung von Gruppierungen liegt im S-Baum Information darüber vor, auf welcher Ebene eine Gruppierung augenblicklich arbeitet und welche Ergebnisse sie liefert. Ein wesentlicher Teil des Konzeptes des S-Baumes liegt in der Möglichkeit, den Knoten Merkmals-Listen zuzuordnen. Dabei entsprechen die Merkmals-Listen der Blätter des S-Baumes den Attributen, die für die entsprechenden Berge im R-Baum gespeichert sind. Für alle anderen Knoten des S-Baumes werden Merkmale aus den Attributen der Berge einer Berg-Gruppe abgeleitet. Die entstehenden Merkmale sind also Gruppen-Merkmale. Solche können sein:

- Intensitäten: z.B. die mittlere, maximale und minimale Intensität der Berge innerhalb einer Berg-Gruppe;
- relative Höhen: z.B. die mittlere, maximale und minimale relative Höhe der Berge innerhalb einer Berg-Gruppe;
- Basisbreiten: z.B. die Gesamtbreite der Berg-Gruppe zusammen mit ihrem linken und rechten Endpunkt und/oder die mittlere, maximale und minimale Breite der Berge einer Berg-Gruppe;
- Texturdeskriptoren: aus den absoluten und relativen Höhen und Basisbreiten der Berge lassen sich Merkmale ableiten, die Texturen innerhalb einer Gruppe beschreiben (vgl. Abschnitt 5.4).

Mit Hilfe dieser Merkmale kann die mehrfache Anwendung von Gruppierungen gesteuert werden. Ändert sich beispielsweise ein Merkmal beim Übergang zur nächsten Ebene sehr stark, dann kann daraus geschlossen werden, daß die gesuchte Region überschritten wurde und verfälschte Werte auftreten. Ferner ist es nicht erstrebenswert, zu große Regionen zu erzeugen, da dies zu trivialen Segmentationen führt. In beiden Fällen kann aufgrund der Merkmale im S-Baum entschieden werden, die Gruppierungen abzubrechen. Durch die Beobachtung von Veränderungen der Werte von Merkmalen

wird das Fensterproblem zumindest teilweise vermieden: wird beim Übergang zur nächsten Ebene eine Region irrtümlich verschmolzen, so werden sich die Werte der Gruppen-Mermale häufig stark ändern. Starke Veränderungen können als Abbruch-Kriterium verwendet werden. Die Rolle des S-Baumes zur Steuerung von Gruppierungen wird in Kapitel 6. noch einmal ausführlich erörtert.

Der S-Baum hat viele Gemeinsamkeiten mit dem Konzept des R-Baumes: auch hier treten Berge als Grundelemente auf, die in Knoten repräsentiert und denen Markmals-Listen zugeordnet sind. Zwischen diesen beiden Baumstrukturen besteht jedoch ein wesentlicher Unterschied: während für ein I-Profil der entsprechende R-Baum in eindeutiger Weise festgelegt ist, hängt die Struktur des S-Baumes von der Art der Gruppierungen und der Häufigkeit ihrer Anwendung ab. Da nacheinander auch verschiedene Gruppierungen eingesetzt werden können, ist es also möglich, für ein I-Profil unterschiedliche S-Bäume zu erzeugen.

5.4 TEXTURANALYSEN

Die Erfassung von Texturen ist eines der wichtigsten und schwierigsten Probleme der Szenenanalyse.

<u>Definition:</u> Eine <u>Textur</u> T wird definiert durch:

(1) Eine Menge $\{ {}^{\tau}E / \tau = 1,\ldots,N \}$ von Texturelementen;

(2) örtliche Verteilungen $\{ {}^{i}V({}^{\tau}E) / i = 1,\ldots,M \}$, für die gilt:

(a) ${}^{i}V({}^{\tau}E)$ legt die Anordnung von Texturelementen fest. Diese kann zufällig oder deterministisch sein.

(b) ${}^{i}V({}^{\tau}E)$ kann als Argument einer örtlichen Verteilung auftreten: ${}^{j}V({}^{i}V({}^{\tau}E))$

Texturen können auf verschiedenen Ebenen erfaßt und beschrieben werden. Die Definition von T (insbesondere (2 b)) ermöglicht die Einführung von Texturebenen:

Definition: Texturebenen TE_i sind durch eine Hierarchie von Texturelementen gegeben, indem sich Elemente ${}^{\tau_i}E$ einer Ebene TE_i aus den Elementen der Ebene TE_{i-1} zusammensetzen, d.h. Elemente ${}^{\tau_i}E$ entprechen Anordnungen von Elementen ${}^{\tau_{i-1}}E$:

(1) $\{{}^{\tau_1}E\} := \{\text{Bildpunkte}\}$

(2) $\{{}^{\tau_i}E\} := \{{}^{i-1}V({}^{\tau_{i-1}}E)\}$

Die Einführung einer Hierarchie von Texturelementen ermöglicht ein rekursives Modell von Texturbeschreibungen. In dieses Modell können verschiedene Texturelemente anderer Ansätze eingeordnet werden:

Ebene TE_1: ${}^{\tau_1}E$ = Bildpunkte

Ebene TE_2: ${}^{\tau_2}E$ = zusammengesetzte Bildpunkte (Grauwertpaare, Kantenelemente, Lauflängen)

Ebene TE_3: ${}^{\tau_3}E$ = zusammengesetzte Kantenelemente (Linienelemente, Flecken)

Ebene TE_4: ${}^{\tau_4}E$ = zusammengesetzte Flecken

Ebene TE_5: ${}^{\tau_5}E$ = ...

In diesem Abschnitt werden Möglichkeiten der Texturerfassung anhand von sechs Texturen untersucht (Abbildung 37). Es handelt sich um Bilder von Kondensatoren, Moos (zweimal), einem Teppich, Kleiderstoff und Widerständen. Diese werden im folgenden stets durch Abkürzungen bezeichnet:

KOND	MOOS I	MOOS II
TEPPICH	KLEID	WID

(in der Reihenfolge von Abbildung 37).

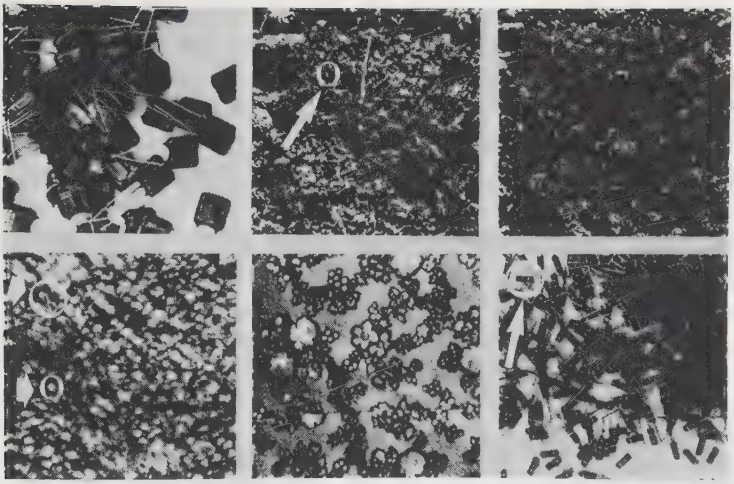


ABBILDUNG 37 : TEXTUREN
(AUSGEWÄHLTE TEXTURELEMENTE SIND DURCH KREISE UND PFEILE GEKENNZEICHNET)

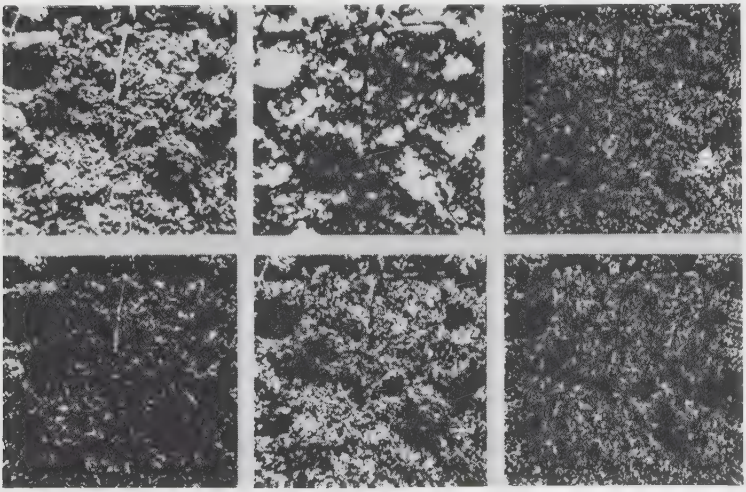


ABBILDUNG 38 : DIE VERTEILUNG VON BILDPUNKTEN IN GRAUWERTSCHEIBEN

(LINKS OBEN : GRAUWERTBILD ; DANACH IM UHRZEIGERSINN : GRAUWERTSCHEIBEN 0-50 , 51-100 , 101-150 , 151-200 , 201-255 ; DIE JEWEILS HELLEN PUNKTE LIEGEN INNERHALB DER GRAUWERTSCHEIBEN)

Die Eigenschaften der Texturelemente legen fest, welche Merkmale auf einer Beschreibungsebene ermittelt werden können. Einzelne Bildpunkte bilden triviale Elemente, da hier nur der Grauwert der Bildpunkte als Attribut zur Verfügung steht. Für die Segmentation können somit nur Schwellwertverfahren (mit Histogrammanalysen) oder Flächenbildung durch Verschmelzen benachbarter Punkte ähnlichen Grauwertes eingesetzt werden. Voraussetzung dieser Verfahren ist eine örtlich (fast) homogene Verteilung diskreter Grauwerte, so daß sich durch Angabe von Grauwertscheiben klar umrissene Häufungsgebiete ergeben. Abbildung 38 zeigt am Beispiel von MOOS I, daß dies nicht zutrifft, sondern daß die Grauwerte der Bildpunkte relativ gleichmäßig in allen Grauwertscheiben auftreten (mit Ausnahme der schwarzen Regionen). Auf dieser Ebene ergeben sich keine guten Beschreibungen.

Betrachten wir noch einmal Abbildung 37. Insbesondere MOOS I, MOOS II und TEPPICH wirken auf den Betrachter als "fleckige Texturen", d.h. sie werden auf TE_3 erfaßt. Wie sehen die Texturelemente dieser Ebene aus? Abbildung 39 zeigt die dreidimensionale Struktur der Texturelemente. Man kann sehen, daß Flecken in der dreidimensionalen Darstellung komplizierte, zusammengesetzte Strukturen aufweisen, in denen die Relationen von Minima und Maxima die Form dieser Strukturen bestimmen (vgl. auch Abb. 11 B). Wenn es gelingt, die Form dieser Texturelemente zu erfassen, dann kann ein mehrdimensionaler Merkmalsvektor auf dieser Texturebene extrahiert werden.

Wie können Texturelemente aus den R-Bäumen gewonnen werden? Die Elemente der Ebene TE_2 sind: Kantenelemente, Grauwertpaare und Lauflängen. Kanten-Elemente zu finden, entspricht dem Problem der Kantendetektion und wurde in Abschnitt 5.2 beschrieben. Grauwertpaare können aus R-Bäumen extrahiert werden, indem alle Punkte auf den Flanken von Bergen ausgewertet werden. Lauflängen entsprechen den Plateaus und sind Grundelemente der verallgemeinerten R-Bäume.

Der Erfassung der örtlichen Verteilungen von Texturelementen wurde bisher in der Literatur relativ wenig Aufmerksamkeit geschenkt. Häufig wird die örtliche Dichte von Texturelementen innerhalb eines Fensters oder

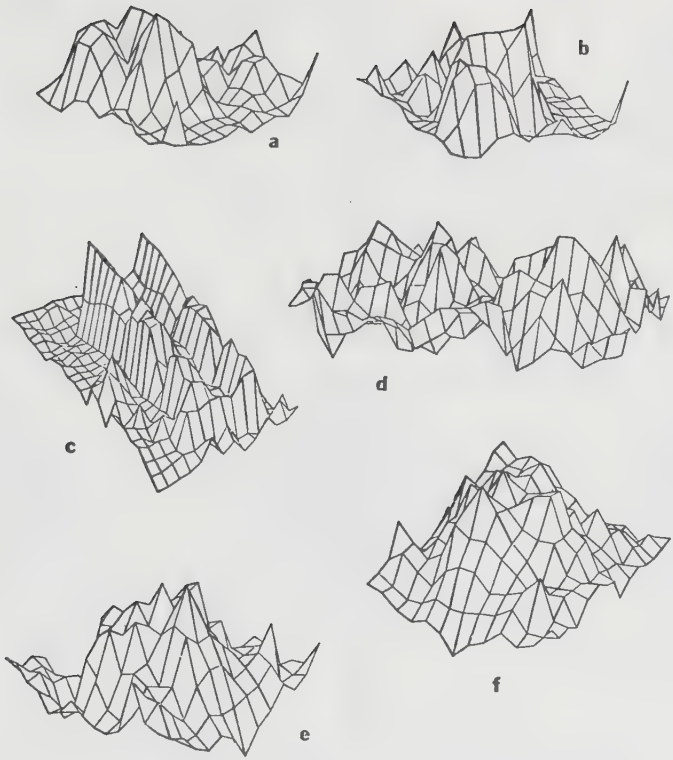


ABBILDUNG 39 : BEISPIELE DER DREIDIMENSIONALEN STRUKTUR VON TEXTURELEMENTEN

(VERGLEICHE AUSGEZEICHNETE ELEMENTE IN ABB. 37)

A,B - ELEMENT AUS MOOS I (UM 60^0 BZW. 120^0 NACH OBEN GEDREHT)

C - ELEMENT AUS WID (INVERTIERT UND UM 60^0 NACH UNTEN GEDREHT)

D - OBERES ELEMENT AUS TEPPICH (UM 20^0 NACH OBEN GEDREHT)

E,F - UNTERES ELEMENT AUS TEPPICH (UM 30^0 NACH UNTEN GEDREHT)

einer Region bestimmt /44, 53/. Bestimmte Konfigurationen von Texturelementen können auch explizit in einem Verfahren enthalten sein. Ein Beispiel sind Auswertungen von Grauwertpaaren, bei denen Bildpunkte verglichen werden, die in einer vorgegebenen Richtung und einem festen Abstand zueinander liegen /14, 33/. Auf die Bestimmung örtlicher Verteilungen wird weiter unten noch einmal kurz eingegangen.

Aus den oben gemachten Beobachtungen ergibt sich, daß Flecken als Elemente der Ebene TE_3 besondere Bedeutung haben. Die Erfassung von Flecken ist in den R-Bäumen besonders einfach, da sie in den I-Profilen Bergen entsprechen und somit Grundelemente der R-Bäume sind. Ein Berg in den R-Bäumen stellt einen Schnitt durch die dreidimensionale Struktur der Flecken dar (Abbildung 40).

Aus den drei Berg-Attributen: Basisbreite, relative und absolute Höhe können nicht-triviale Textur-Merkmale gewonnen werden. Ein einfaches Experiment zeigt die Vorteile einer Erfassung von Texturen auf dieser Ebene.

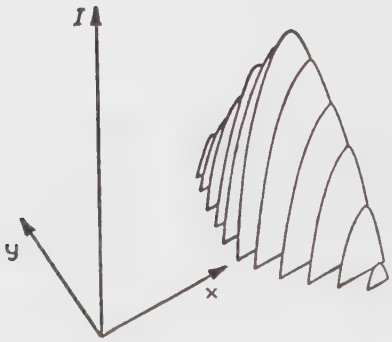


ABBILDUNG 40 : SCHNITTE DURCH DIE DREIDIMENSIONALE STRUKTUR EINES TEXTURELEMENTES

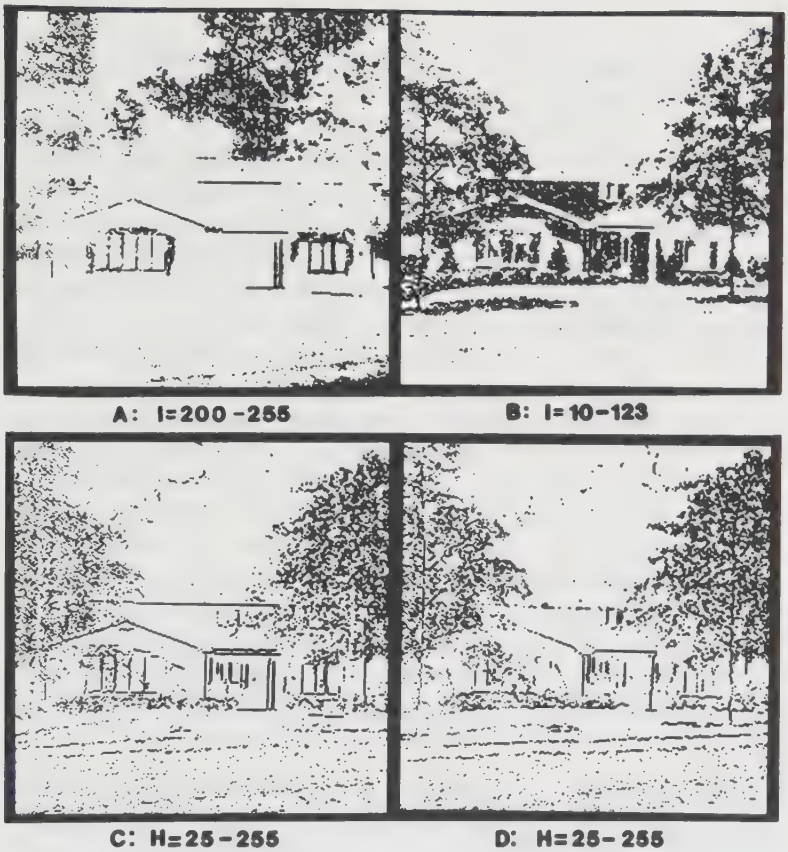


ABBILDUNG 41 : DIE ÖRTLICHE VERTEILUNG VON MINIMA UND MAXIMA FÜR VERSCHIEDENE WERTEBEREICHE VON BERG-ATTRIBUTEN

A - MAXIMA B - MINIMA ATTRIBUT : ABSOLUTE HÖHE I
C - MAXIMA D - MINIMA ATTRIBUT : RELATIVE HÖHE H

In Abbildung 41 sind die Ergebnisse von Schwellwertoperationen bezüglich der Attribute: relative und absolute Höhe der Berge dargestellt.
Wie in Abschnitt 5.3.1 (Abbildung 31) ist zu beachten, ob helle oder dunkle Flecken vorliegen. Während die hellen Flecken direkt den Bergen entsprechen, treten dunkle Flecken als Täler auf. Um diese zu erfassen, wird das Bild invertiert, (Grauwert 0 = Weiß, Grauwert 255 = Schwarz),

so daß die dunklen Flecken zu Bergen im entstehenden I-Profil werden. Abbildung 41 A zeigt die örtliche Verteilung der Maxima, deren zugehörige Berge absolute Höhen innerhalb des Wertebereiches 200 - 255 haben. Abbildung 41 B zeigt die entsprechende Verteilung für Minima. Relative Höhen sind ein Maß für den lokalen Kontrast eines Berges. Viele Texturen sind nicht durch ihren mittleren Grauwert charakterisiert, sondern durch die Varianz des Grauwertes innerhalb einer Region. Dies entspricht ungefähr dem lokalen Kontrast und damit der relativen Höhe von Bergen. Abbildungen 41 C, D zeigen die örtliche Verteilung der Maxima und Minima bezüglich eines Wertebereiches der relativen Höhen von 25 - 255. Aus Abbildung 41 ersieht man, daß viele Regionen des Grauwertbildes in einer der vier Darstellungen als ausgeprägte Häufungsgebiete auftreten. In Abbildung 41 C, D wird insbesondere die Wirksamkeit der Verarbeitung relativer Höhen deutlich: in beiden Fällen sind die Bäume rechts und links des Hauses vom Dach getrennt worden. Dieses Ergebnis ist durch Schwellwertoperationen aufgrund der Grauwerte nicht zu erzielen, da viele Bildpunkte des Daches denselben Grauwert wie Bildpunkte der Bäume haben. Selbst Konturen treten in den Ergebnissen von Abbildung 41 als geschlossene Linien auf. Ein Problem einer derartigen Verarbeitung liegt in der automatischen Ermittlung geeigneter Wertebereiche für die Attribute. Ebenso fehlen Verfahren zur Begrenzung der auftretenden Häufungsgebiete.

Berge bieten neben ihren drei Attributen einen weiteren Vorteil. Texturelemente der Ebenen TE_1 und TE_2 können nur durch die Relation "Konkatenation" verknüpft werden. Berge treten jedoch nicht nur nebeneinander, sondern auch ineinander verschachtelt auf. Diese Verschachtelung wird durch die Struktur der R-Bäume ausgedrückt. Der Übergang von der Texturebene TE_3 zu höheren Ebenen TE_4, TE_5,.... kann durch die Gruppierungen von Bergen erfolgen z.B. die Splitter- oder Stamm-Gruppierung (vgl. Abschnitt 5.3.1). Der Zusammenhang zwischen den verschiedenen Texturebenen wird durch den S-Baum festgehalten (vgl. Abschnitt 5.3.2). Jedem Knoten im S-Baum, der eine Berg-Gruppe repräsentiert, kann ein <u>Texturdeskriptor</u> in Form eines Merkmalvektors zugeordnet werden. Im folgenden wird gezeigt, wie geeignete Merkmale aus den Berg-Attributen abgeleitet werden können.

Die gebräuchlichste Technik zur Bestimmung von Texturdeskriptoren besteht darin, Histogramme von Attributen auszuwerten. Wenn Berge als Texturelemente verwendet werden, stehen die drei Attribute: Basisbreite, absolute und relative Höhe zur Verfügung. Andere Ansätze verwenden ähnliche Attribute, in der Regel jedoch nur jeweils eines der drei. Breiten werden beispielsweise durch einen verallgemeinerten Laplace-Operator ermittelt /77/. Intensitäten werden häufig für Methoden der Flächenbildung durch Schwellwertverfahren und Histogrammanalysen ausgewertet. Experimente, in denen der Kontrast von Texturelementen ausgewertet wird, haben die Brauchbarkeit dieses Attributes gezeigt /10, 48/. Allerdings wurden in keinem Ansatz alle drei Attribute gleichzeitig verwendet. Dies ist durch Einsatz der R-Bäume möglich, da hier die entsprechenden Attribute gespeichert sind.

Bei der Berechnung von Histogrammen müssen die Texturebenen vorher voneinander getrennt werden, da sonst Texturelemente verschiedener Ebenen zu den gleichen Histogrammelementen beitragen und zu verfälschten Werten führen. Dadurch werden Häufungsgebiete in den Histogrammen verzerrt oder verlagert,und man erhält nicht die richtigen Deskriptoren. Eine direkt zugängliche Texturebene TE_3 ist in den Blättern der R-Bäume enthalten. Um zu einer höheren Texturebene TE_4 zu gelangen, wurde die Stamm-Gruppierung angewandt und die Attribute der entstehenden Berge verwertet.

Zur Erprobung der Brauchbarkeit der Berg-Attribute für die Merkmalextraktion wurden mit den Texturen aus Abb. 37 Experimente durchgeführt.
Im Verlauf der Experimente ergaben sich Beobachtungen, die die Art der Auswertung der drei Attribute festlegten. Es zeigte sich, daß die Erfassung der Attribute in Zeilenrichtung nur zu bedingt brauchbaren Texturdeskriptoren führt. Der Grund hierfür wird aus Abbdilung 40 deutlich.
Da Berge Schnitte durch die eigentlichen Texturelemente, die Flecken, sind, ergeben sich zu hohe Werte für kleine Basisbreiten und niedrige absolute und relative Höhen. Aus diesem Grund werden in der Auswertung nur solche Berge berücksichtigt, die sowohl in Zeilen- als auch in Spalten-Richtung auf einem Maximum liegen. Das heißt, es werden R-Bäume in Zeilen- und Spalten-Richtung konstruiert und nur diejenigen Berge ausgewertet, die in beiden R-Bäumen auftreten. Als relative Höhe wird die

kleinere der relativen Höhen einer der beiden Richtungen gewählt. Eine weitere Beobachtung zeigt, daß bei den untersuchten Texturen keine wesentlichen Unterschiede der Basisbreiten in Zeilen- oder Spalten-Richtung auftreten, so daß es genügt, die Basisbreite in einer Richtung (zeilenweise) zu erfassen. Es ist klar, daß bei Texturen, in denen Flecken ausgeprägter Richtung vorherrschen, beide Breiten berücksichtigt werden müssen.

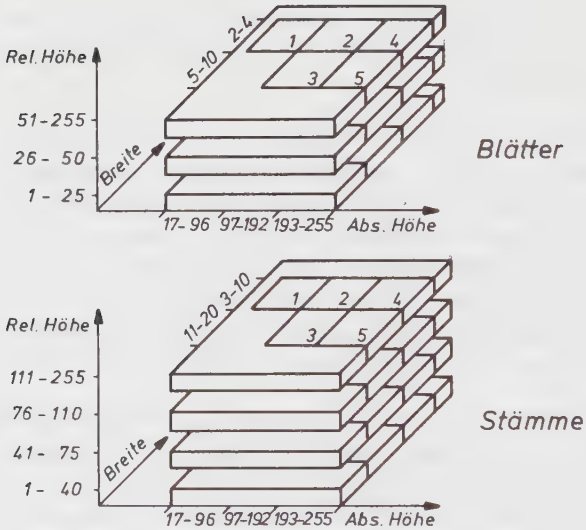


ABBILDUNG 42 : DIE AUSWERTUNG DREIDIMENSIONALER HISTOGRAMME DURCH UNTERTEILUNG IN WERTEBEREICHE

EBENE TE_3 (BLÄTTER) : BREITE = 2- 4 , 5-1o
REL. HÖHE = 1-25 , 26-5o , 51-255
ABS. HÖHE = 17-96 , 97-192, 193-255

EBENE TE_4 (STÄMME) : BREITE = 3-1o , 11-2o
REL. HÖHE = 1-4o , 41-75 , 76-11o , 111-255
ABS. HÖHE = 17-96 , 97-192, 193-255

Aus diesen Attributen werden dreidimensionale Histogramme berechnet. Um diese Histogramme auswerten zu können, wird auf jeder Achse des Histogrammes eine Unterteilung in verschiedene Wertebereiche vorgenommen (siehe Abbildung 42). Durch die Unterteilung der X-Achse (Absolute Höhen) in drei Wertebereiche unterscheidet man zwischen dunklen/grauen/hellen Flecken. Die Unterteilung der Y-Achse (Basisbreiten) liefert schmale/breite Berge und die der Z-Achse (Rel. Höhen) ermöglicht eine Unterscheidung von niederen/mittleren/hohen Bergen. Für die untersuchten Texturen werden auf Ebene TE_3 15 Merkmale ausgewählt. Diese werden gewonnen, indem Histogrammelemente, die innerhalb eines Wertebereich-Tripels liegen, aufsummiert werden. So entspricht z.B. ein Merkmal (vgl.Abbildung 42) Bergen mit Basisbreiten zwischen 2 - 4, relativen Höhen zwischen 51 - 255 und absoluten Höhen zwischen 17 - 96. Für Texturen der Ebene TE_4 erweist sich eine Unterteilung der Z-Achse in vier Wertebereiche als notwendig, wobei 20 Merkmale für die Beschreibung ausgewählt werden.

Für beide Texturebenen zusammen liegen somit 35 Merkmale vor, die für die Texturen aus Abb. 37 berechnet wurden. In der Auswertung ergab sich, daß vier der Merkmale in allen sechs Texturen gleich Null sind, so daß diese Merkmale in der Darstellung der Ergebnisse weggelassen werden. Die verbleibenden 31 Merkmale werden in einem Merkmalsvektor erfaßt, wobei die Reihenfolge der Merkmale nach aufsteigender relativer Höhe erfolgt. Abbildung 43 enthält eine graphische Darstellung der Merkmalsvektoren, wobei eine Normierung für jedes Merkmal durchgeführt wurde.

Wenn diese Merkmalsvektoren gute Texturdeskriptoren darstellen sollen, muß es möglich sein, Zusammenhänge zwischen Merkmalen mit hohen Werten und visuellen Charakteristika der Texturen herzustellen. Dies ist in der Tat möglich:

KOND - TE_3: die vielen hellen Berge (195 - 255) mit niederen relativen Höhen deuten auf helle homogene Regionen; viele graue Berge (97 - 192) mittlerer relativer Höhe und Breiten von 5 - 10 lassen auf graue Flecken auf schwarzem Grund schließen.

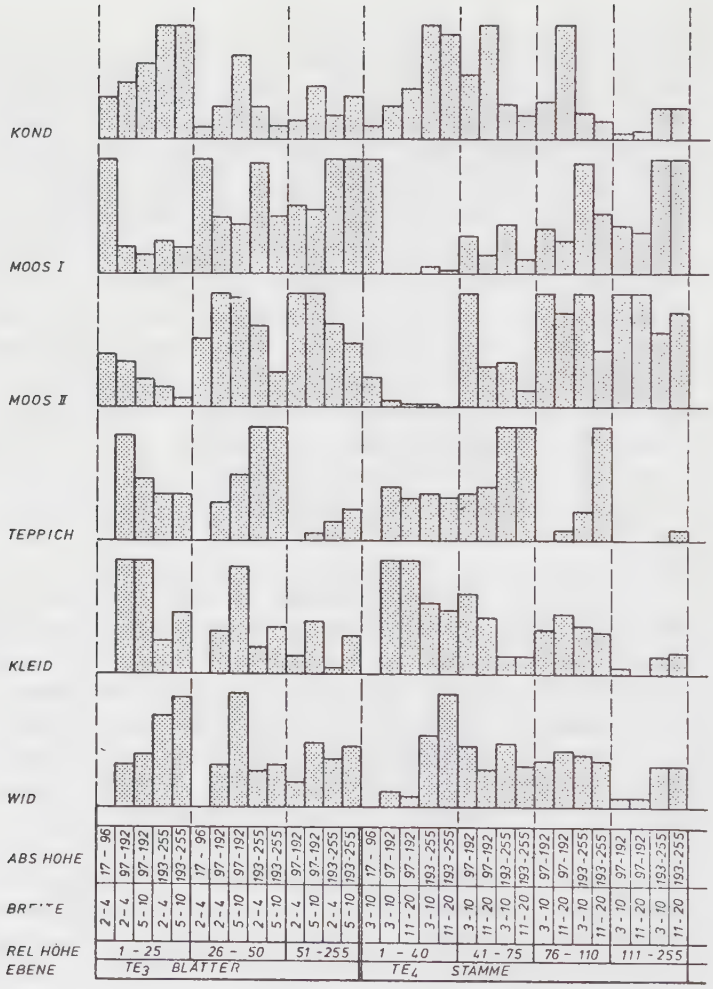


ABBILDUNG 43 : NORMIERTE MERKMALSVEKTOREN DER TEXTUREN AUS ABB. 37

KOND - TE_4: ähnlich wie TE_3. Die Tatsache, daß nur wenige Stämme mit hohen relativen Höhen auftreten, läßt auf niederen Kontrast im Bild oder auf große homogene Regionen schließen.

MOOS I - TE_3: Niedere dunkle Berge deuten auf homogene schwarze Regionen; graue und helle Berge mit hohen relativen Höhen sind ein Hinweis auf eine harte Textur.

MOOS I - TE_4: Viele helle Berge großer relativer Höhe zeigen eine kontrastreiche Textur an.

MOOS II - TE_3: Im Vergleich zu MOOS I hat diese Textur weniger schwarze Regionen; statt dessen findet eine Verschiebung auf graue Berge statt.

MOOS II - TE_4: Auch auf dieser Ebene liegt eine Verschiebung zu grauen Werten vor.

TEPPICH - TE_3: Viele graue Berge mit niederem Kontrast sowie das Fehlen von Bergen mit hohem Kontrast weisen auf eine weiche Textur hin.

TEPPICH - TE_4: Auch hier sind fehlender Kontrast und viele breite Berge mit niederen relativen Höhen kennzeichnend.

KLEID - TE_3: Dunkle Berge niederen Kontrastes erfassen die schwarzen Regionen im Muster; graue Berge erfassen den Untergrund,und Berge hohen Kontrastes stammen von den hellen Flecken der Blumen.

KLEID - TE_4: Die Anzahl der Berge fällt mit steigendem Kontrast: ein Zeichen für große Regionen einer schwachen Textur.

WID - TE_3: Viele helle Berge mit geringen relativen Höhen deuten auf helle homogene Regionen.

WID - TE_4: Auch hier weist ein hoher Anteil heller Berge geringen Kontrastes auf helle homogene Regionen.

Die Merkmale erfassen also wesentliche Charakteristika der Texturen. Diese Merkmalsvektoren können zur Klassifikation von Texturen eingesetzt werden. Dies war jedoch nicht Ziel der Experimente und soll hier nur exemplarisch an zwei Beispielen gezeigt werden. Ein Vergleich zwischen zwei Texturen kann auf einfache Weise durch die Differenz ihrer normierten Merkmalsvektoren vorgenommen werden. Da die Differenzen für jedes Merkmal ermittelt werden, erhält man wieder einen Vektor. Abbildung 44 zeigt zwei dieser Differenzvektoren, die das Ergebnis

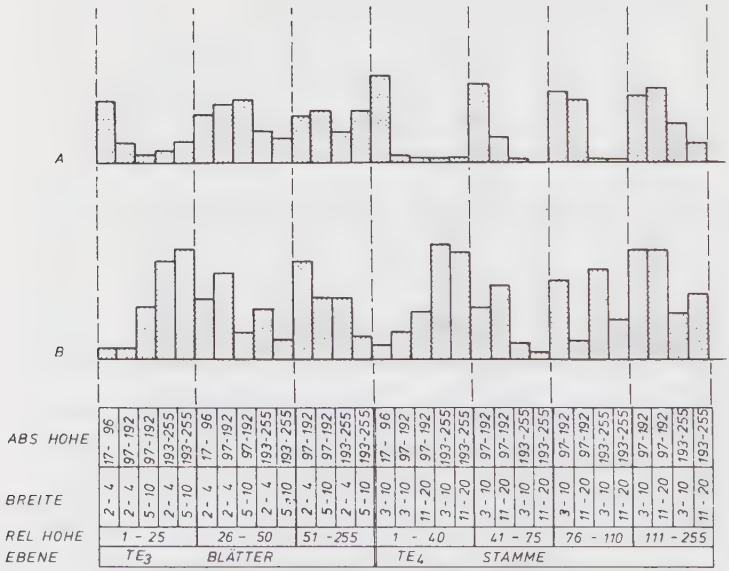


ABBILDUNG 44 : DIFFERENZVEKTOREN FÜR DEN VERGLEICH ZWISCHEN TEXTUREN

A - MOOS I/MOOS II B - MOOS I/KOND

eines Vergleiches zwischen MOOS I/MOOS II bzw. MOOS I/KOND darstellen. Die beiden Moos-Texturen erweisen sich in Bezug auf TE_4 in einigen Merkmalen als sehr ähnlich, jedoch nicht in allen. Typischerweise treten die Unterschiede vor allem für kleine absolute Höhen (17 - 96) auf. Auf TE_3 sind die Unterschiede breiter gestreut, jedoch quantitativ geringer. Hier treten Unterschiede vorwiegend für mittleren und hohen Kontrast (Rel. Höhe 26 - 50, 51 - 255) auf. Die Unterschiede zwischen den beiden Texturen sind durch eine Kontrastverschiebung von Elementen mittleren und hohen Kontrastes auf TE_3 und durch eine Verschiebung zu geringeren Intensitäten auf TE_4 (z.T. auch auf TE_3) gekennzeichnet. Dies war zu erwarten, da MOOS I aus MOOS II erzeugt wurde, indem die Grauwerte aus MOOS II innerhalb von 2 x 2 Fenstern durch den durchschnittlichen Grauwert ersetzt wurden. Aufgrund dieser Operation ist für MOOS I schwächerer Kontrast und geringere Intensität zu erwarten. Die Unterschiede, die sich aus dem Differenzvektor ergeben, entsprechen den Unterschieden, die bei Betrachtungen der beiden Texturen wahrgenommen werden.

Im Gegensatz zu den systematischen Unterschieden zwischen den beiden Moos-Texturen stehen die Unterschiede zwischen MOOS I und KOND (Abb. 44 B). Hier sind die Unterschiede in vielen Merkmalen groß, und es läßt sich keine systematische Verschiebung ablesen. Dies entspricht auch den visuellen Eindrücken.

Merkmale,wie die oben eingeführten, erfassen charakteristische Werte von Texturen, die örtlichen Verteilungen ${}^{i}V({}^{\tau}E)$ der Texturelemente bleiben dabei jedoch unberücksichtigt. Die Erfassung dieser Information ist möglich, indem die Koordinaten der Berge ermittelt werden, die in einem Merkmal auftreten. Abbildung 45 enthält Darstellungen von Beispielen, wobei jeder Berg durch eine Raute dargestellt ist, deren Ausmaße den Basisbreiten entsprechen. Insbesondere die Beispiele aus Abbildung 45 A, B machen deutlich, daß zusätzliche Information in der örtlichen Verteilung vorhanden ist: bei fast gleicher Anzahl von Bergen (233 bzw. 240) treten die Berge in Abbildung 45 A in Häufungsgebieten auf, während sie in Abbildung 45 B fast gleichmäßig verteilt sind. Um auch Informationen über örtliche Verteilungen in der Texturbeschreibung berücksich-

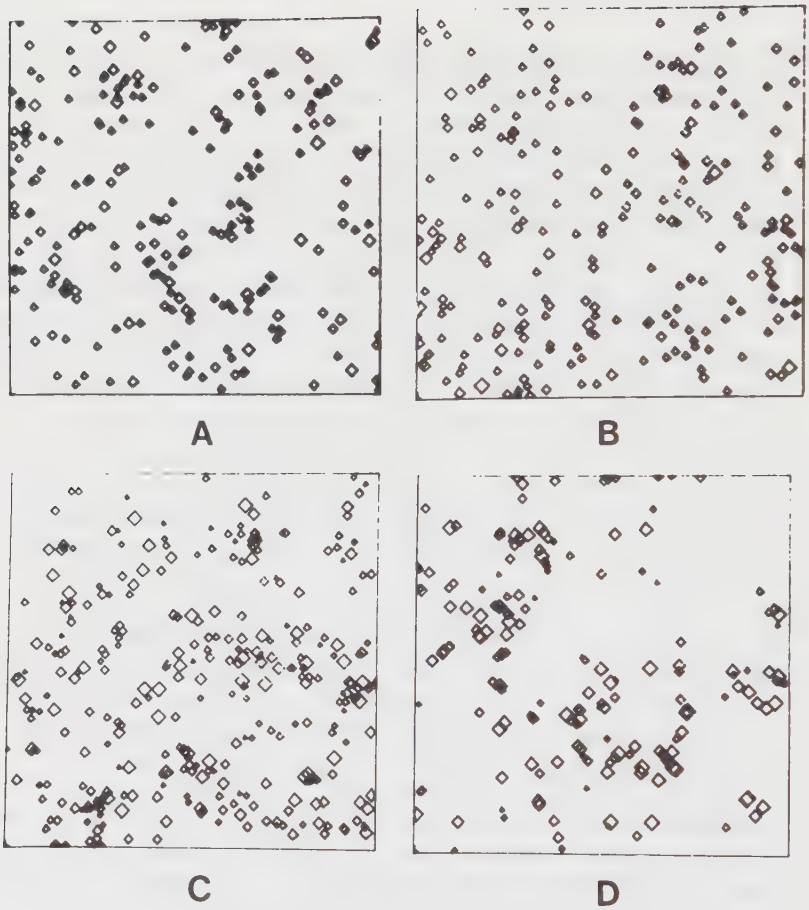
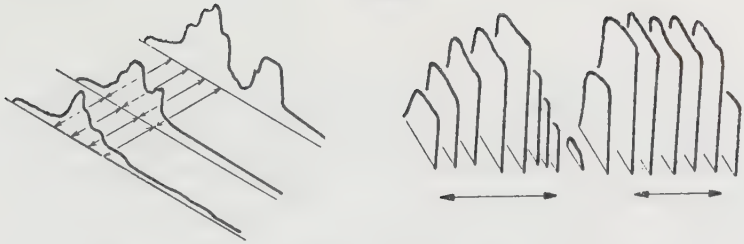


ABBILDUNG 45 : DIE ÖRTLICHE VERTEILUNG VON BERGEN

A - KLEID-TE_3 : BREITE=5-1o ; REL.HÖHE=26-5o ; ABS.HÖHE=97-192
B - MOOS I-TE_3 : BREITE=5-1o ; REL.HÖHE=26-5o ; ABS.HÖHE=97-192
C - MOOS I-TE_4 : BREITE=3-1o ; REL.HÖHE=41-75 ; ABS.HÖHE=97-192
D - KOND-TE_4 : BREITE=3-1o ; REL.HÖHE=41-75 ; ABS.HÖHE=97-192



A - VERKNÜPFUNG VON BERGEN B - ABBRUCH DER VERKNÜPFUNG BEI ERREICHEN EINES TALS

ABBILDUNG 46 : PROFILVERFOLGUNG

tigen zu können, besteht die Möglichkeit, für jedes der Merkmale Kenngrößen über seine Verteilung zu ermitteln. Ein Beispiel für ein solches Maß ist die Berechnung der Abstände der k-nächsten Nachbarn eines jeden Elementes einer Verteilung. Derartige Informationen können anschließend verwendet werden, um Grenzlinien von Texturen zu finden.

5.5 DIE VERARBEITUNG QUER ZU DEN ZEILEN: PROFILVERFOLGUNG

Der Ansatz der R-Bäume stellt eine zeilenweise Beschreibung von Grauwertbildern zur Verfügung und muß durch zusätzliche Maßnahmen die zweidimensionale Bildinformation wiedergewinnen. Hierzu wurden in den vorangegangenen Abschnitten wiederholt die R-Bäume sowohl in Zeilen- als auch Spalten-Richtung konstruiert. Eine Alternative bietet hier eine Verarbeitung quer zu den Zeilen, bei der Teilstrukturen der I-Profile über mehrere Zeilen hinweg verfolgt werden. Dieser Vorgang soll "Profilverfolgung" genannt werden.

Die Verarbeitung in dieser Richtung findet für I-Ereignisse, die sich aus einer Auswertung der R-Bäume ergeben haben, den Anschluß an I-Ereignisse der Nachbarzeilen und verknüpft diese Ereignisse miteinander. Solche I-Ereignisse können Kantenelemente, Berge, Plateaus, Berg-Gruppen oder andere Teilstrukturen der I-Profile sein. Im weiteren wird exemplarisch die Verknüpfung von Bergen erörtert.

Die Aufgabe der Profilverfolgung besteht somit darin, für einen Berg in einer Zeile einen Berg in den Nachbarzeilen zu finden, der eine Fortset-

zung darstellt (Abbildung 46 A). Die Profilverfolgung gliedert sich in drei Schritte: (1) Finden von Start-Elementen; (2) Bestimmung von Anschluß-Elementen und (3) Abbruch der Verknüpfung, wenn kein sinnvoller Anschluß mehr möglich ist.

Die Verfolgung von Bergen bricht ab, wenn ein Tal erreicht ist, so daß auch quer zu den Zeilen Berge und Plateaus entstehen (Abbildung 46 B).

Zwischen der Verarbeitung in Zeilenrichtung und der Verarbeitung quer zu den Zeilen besteht ein wichtiger Unterschied: im ersten Fall werden Bildpunkte verarbeitet - und zwar jeweils alle Punkte einer Zeile; im zweiten Fall werden Berge erfaßt - nicht alle, sondern jeweils wenige, die für die Analyse wichtig sind. Bei der Profilverfolgung werden also nicht alle möglichen Verknüpfungen vorgenommen, sondern es findet eine zielgerichtete Verarbeitung statt. Die Verknüpfungen zwischen Elementen benachbarter Zeilen erfolgt durch Verweise in den entsprechenden Knoten der R-Bäume (Abbildung 47).

Das zielgerichtete Verhalten schlägt sich insbesondere in der Auswahl der Startpunkte und im Abbruch der Profilverfolgung nieder. Während der Auswertung in Zeilenrichtung kann bereits eine Auswahl getroffen werden, welche Berge für eine Verfolgung interessant sind. Durch Markieren solcher Berge als "auffällig", liegt nach der Auswertung der R-Bäume eine Liste von Kandidaten für die Startpunkte der Profilverfolgung vor. Die Auswahl auffälliger Berge kann dabei nach rein syntaktischen Gesichtspunkten erfolgen, z.B. indem Berge großer Basisbreite und hohen Kontrastes markiert werden. Die Auswahl kann aber auch mit Hilfe von A-priori-Wissen erfolgen, wenn z.B. die Struktur eines I-Profiles über einem Objekt bekannt ist.

Ein wesentlicher Teil des Konzeptes der Profilverfolgung liegt in der Möglichkeit, die Verknüpfungen in hierarchischer Weise vorzunehmen, indem man mit Knoten nahe der Wurzel beginnt. Diese Knoten entsprechen großen Regionen und sind deshalb sicher zu verfolgen. Erst nachdem die sicher zu verfolgenden Regionen verknüpft sind, geht man zu kleineren Regionen über. Auf diese Weise werden Entscheidungen durch bereits getroffene

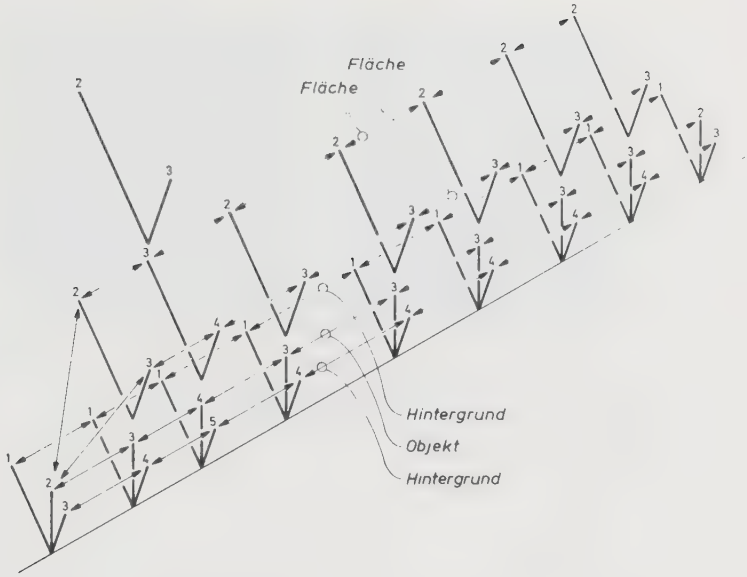


ABBILDUNG 47 : DIE VERKÜPFUNG VON BERGEN IN R-BÄUMEN

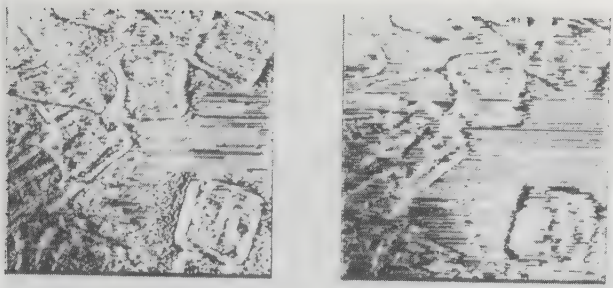
unterstützt, und es wird verhindert, daß die Profilverfolgung sich über Grenzlinien hinaus verirrt.

Nach Auswahl von Startpunkten müssen Berge gefunden werden, die einen Anschluß ermöglichen. Kandidaten für eine Verknüpfung sind Berge, deren Basis mit der des Start-Berges gemeinsame Punkte hat. Ein Beispiel dafür, wie die Verknüpfung von Bergen zustandekommt, bietet ein einfacher Algorithmus, der Zeiger zwischen Bergen und Tälern benachbarter Zeilen zieht. Der Algorithmus beruht darauf, für jeden Berg einer Zeile den nächstgelegenen Berg der Nachbarzeile zu finden. Dieser Vorgang wird rückwärts für die Nachbarzeile wiederholt. Die Berge der beiden Zeilen, die sich gegenseitig am nächsten sind, werden durch einen Zeiger verbunden. Mit Tälern wird analog verfahren. Abbildung 48 zeigt ein Beispiel. Erfahrungen mit diesem Algorithmus zeigen, daß es günstig ist, mehrere Berge gleichzeitig zu verfolgen. Die Berücksichtigung mehrerer Berge

führt dazu, daß falsche Berg-Verknüpfungen sich gegenseitig blockieren, so daß sich die Profilverfolgung nicht seitlich verirrt.

Abbildung 49 zeigt eine andere Darstellung der Ergebnisse dieses Algorithmus. Ziel dieser Darstellung ist es, die Regionen zu zeigen, die durch Verknüpfung der Berge entstehen. Anstelle der vertikalen Zeiger (vgl. Abbildung 48) wird zu diesem Zweck die Projektion der Basis jedes Berges, für den ein Anschluß in der Nachbarzeile besteht, als heller horizontaler Strich dargestellt. Innerhalb dieser Linie kennzeichnet ein heller Punkt die Lage des Maximums. Da die Verknüpfung unabhängig von den absoluten Höhen der Berge stattfindet, werden sowohl Berge im Hintergrund als auch die Berge, die auf Objekten liegen, miteinander verknüpft. Die Wirkung der Verknüpfung läßt sich am besten an den dunklen Stellen in Abb. 49 ablesen, da dort keine Verknüpfung erfolgt . Die dunklen Stellen fallen also mit den Grenzlinien von Regionen zusammen.

ABBILDUNG 48 : DIE VERKNÜPFUNG VON BERGEN UND TÄLERN IN INTENSITÄTS-PROFILEN



A B

ABBILDUNG 49 : DIE VERKNÜPFUNG VON BERGEN DES GRAUWERTBILDES AUS ABB. 25 A

A - VERKNÜPFUNGSEBENE : BLÄTTER B - VERKNÜPFUNGSEBENE : STÄMME

Abbildung 49 A gibt die Verknüpfung von Blättern der R-Bäume wieder. Es werden viele Berge verknüpft, die jedoch nur kleine Regionen bilden. Dies ist zu erwarten, da die Blätter Regionen kleiner Ausmaße entsprechen. Es werden zwar nicht alle Berge verknüpft, aber auch nicht Grenzlinien von Regionen überschritten. Abb. 49 B zeigt die Verknüpfung von Stämmen. Hier haben Berge eine breitere Basis, so daß die entstehenden Regionen besser ausgefüllt sind. Auch hier überschreitet der Algorithmus nur selten die Grenzen von Regionen. Für die hier gezeigten Verknüpfungen wurden lediglich die Abstände von Bergen in Nachbarzeilen benutzt. Es sind jedoch Verfeinerungen denkbar, bei denen auch von Berg-Attributen Gebrauch gemacht wird, um über Verknüpfungen zu entscheiden.

Ein zielgerichtetes Verhalten der Profilverfolgung setzt voraus, daß die Ergebnisse der Verknüpfungen laufend ausgewertet werden, um über Abbruch oder Fortsetzung entscheiden zu können. Es ist also ein "Buchhalter"-Modul notwendig, in dem Information über die entstehenden Regionen gespeichert wird. Als Information stehen zur Verfügung:

(1) die Fläche der Regionen aus den Basisbreiten der verknüpften Berge (vgl. Abb. 49);

(2) Lage und Verlauf der Konturen aus den x^L und x^R der verknüpften Berge (vgl. Abb. 49);

(3) Texturinformationen, die aus den Berg-Attributen berechnet werden (vgl. Abschnitt 5.4).

Mit Hilfe solcher Informationen kann nicht nur über den Abbruch der Profilverfolgung entschieden werden, sondern auch darüber, ob ein Übergang auf eine feinere Verfolgung notwendig ist, bei der Knoten verknüpft werden, die im R-Baum in der Nähe der Blätter liegen. Eine Implementierung dieser Maßnahmen ist jedoch bisher nicht erfolgt.

6. EIN MODELL NIEDERER PROZESSE FÜR SZENENANALYSEN

In den vorangegangenen Kapiteln wurde beschrieben, wie ein Grauwertbild in eine P-Repräsentation überführt und diese durch verschiedene Techniken ausgewertet wird. Relational-Bäume liefern die symbolische P-Repräsentation; ihre Auswertung erfolgt durch die Detektion von I-Ereignissen, Gruppierungen und Profilverfolgung. In Abschnitt 3.2 wurde die S-Repräsentation als Schnittstelle zwischen niederen und höheren Prozessen definiert. Ein Modell niederer Prozesse muß also das Zusammenspiel der verschiedenen Auswertungs-Moduln und die Art der entstehenden S-Repräsentation erklären. Das hier vorliegende Modell beschreibt Vorgänge der Segmentation und Merkmalsextraktion und erlaubt, neben den eigenen Ansätzen auch Verfahren aus der Literatur in das Konzept zu integrieren. Gleichzeitig wird erläutert, wie der vorliegende Ansatz weiterentwickelt werden kann.

6.1 DIE KONTROLL-STRUKTUR DES MODELLS

Abbildung 50 enthält den Aufbau des Modelles. Es soll hier erklärt werden, welche Aufgaben die einzelnen Moduln haben, in welcher Weise Informationen zwischen den Moduln fliessen und wie Entscheidungen getroffen werden. Im Einzelnen sind zu beachten:

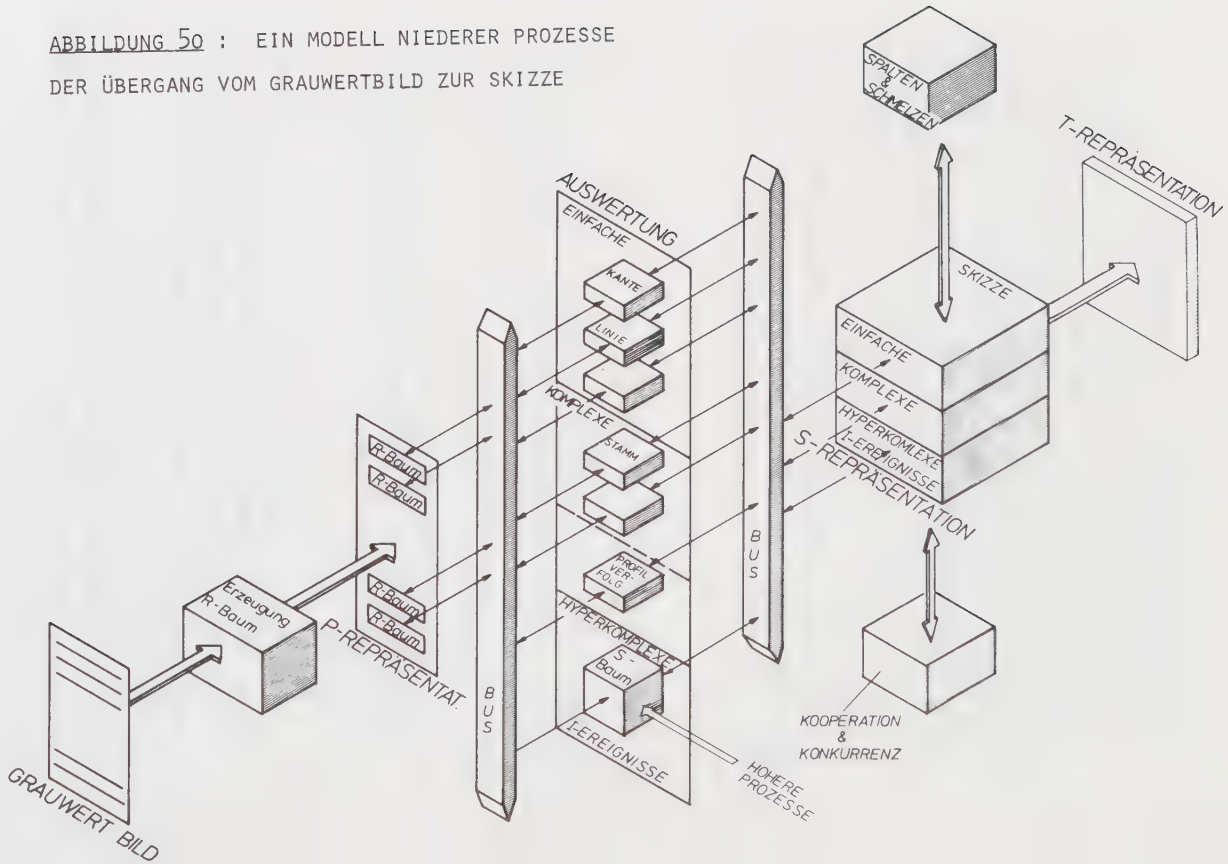
- Der deterministische Übergang vom Grauwertbild zur P-Repräsentation. Diese besteht aus einer Folge von R-Bäumen, die eine hierarchische Strukturbeschreibung der I-Profile aller Bildzeilen liefern.

- Die Auswertungs-Moduln, die auf die P-Repräsentation zugreifen und diese nach verschiedenen Kriterien verarbeiten. Entsprechend der Einteilung in einfache, komplexe und hyperkomplexe I-Ereignisse wird eine Vielzahl solcher Moduln eingesetzt. Eine besondere Rolle spielt hierbei das Struktur-Baum-Modul (S-Baum-Modul).

- Die Darstellung des Informationsflusses; die einzelnen Moduln sind - wie auch bei Mikroprozessoren - über ein Bus-System verbunden ,

ABBILDUNG 5o : EIN MODELL NIEDERER PROZESSE

DER ÜBERGANG VOM GRAUWERTBILD ZUR SKIZZE



das allen Moduln gemeinsam zur Verfügung steht.

- Die Speicherung der Auswertungs-Ergebnisse in der S-Repräsentation. Diese wird im Anschluss an andere Arbeiten / 7,44/ "Skizze" genannt und enthält mehrere Ebenen, in denen einfache, komplexe und hyperkomplexe I-Ereignisse gespeichert sind.

- Die Bereinigung der Skizze durch Kooperation und Konkurrenz von Ergebnissen bis ein konsistentes Ergebnis vorliegt.

- Die Auswahl einer Beschreibungsebene durch Verfahren des Spaltens und Schmelzens.

DER ÜBERGANG VON GRAUWERTBILD ZUR P-REPRÄSENTATION

Der Übergang vom Grauwertbild zur P-Repräsentation verläuft im Modell deterministisch, die Struktur des Grauwertgebirges (in Form von I-Profilen) bleibt weitgehend erhalten. Dieser Übergang stellt eine Aufbereitungsphase dar, in der eine Repräsentation erzeugt wird, die:

- globale und lokale Analysen ermöglicht;
- die Erfassung von Konturen und Regionen unterstützt;
- eine Beschreibung von I-Ereignissen (Kanten, Linien, Texturen) liefert.

Ein wesentlicher Vorgang bei diesem Übergang ist das Sammeln von Kontext-Information, die die anschließende Verarbeitung erleichtert. Der Gebrauch von Relational-Bäume zeigt exemplarisch eine geeignete P-Repräsentation; es gibt jedoch auch andere mögliche Repräsentationen.

DIE AUSWERTUNG DER P-REPRÄSENTATION

Die P-Repräsentation wird nach verschiedenen Kriterien verarbeitet. Dies geschieht durch Verfahren, die die P-Repräsentation in Intensitäts-Ereignisse zerlegen. Bei diesem Verarbeitungsschritt sind Fehler und Widersprüche der Ergebnisse unvermeidlich. Dies führt dazu, daß für Bildpunkte oder Regionen mehrere, eventuell widersprüchliche Auswertungs -

Ergebnisse in die Skizze eingetragen werden. Es ist Aufgabe der Bereinigung der Skizze, derartige Widersprüche aufzulösen (siehe weiter unten).

Auswertungs-Techniken wurden in Kapitel 5. entwickelt: einfache I-Ereignisse wie Kanten- oder Linien- Elemente können direkt den R-Bäumen entnommen werden. Komplexe I-Ereignisse werden durch Gruppierungen gewonnen, die auch rekursiv angewandt werden können. Zusätzlich besteht die Möglichkeit, verschiedene Gruppierungen gemischt nacheinander einzusetzen. Insbesondere beim Einsatz verschiedener Gruppierungen ist zu erwarten, daß widersprüchliche Ergebnisse auftreten, die später bereinigt werden müssen. Für die Bereinigung der Skizze ist es wichtig, eine Abschätzung für die "Stärke" eines Auswertungs-Ergebnisses vornehmen zu können. Auch kann nicht erwartet werden, daß eine Gruppierung überall zu sinnvollen Ergebnissen führt, d.h. es muß möglich sein, falsche Ergebnisse zu entfernen. Damit ergeben sich folgende Fragen:

- Wie kann die Stärke von I-Ereignissen bestimmt werden?
- Nach welchen Kriterien werden Gruppierungen mehrfach angewandt?
- Wann werden Gruppierungen abgebrochen?

Die Bestimmung der Stärke von Ergebnissen läßt sich durch eine Erweiterung der Relational-Bäume erreichen, die hier lediglich angedeutet wird. Diese Erweiterung besteht in einer Gewichtung der Knoten, die für jeden Knoten besagt, wie "stark" er ist. Die Stärke von Knoten wird bestimmt durch die Lage des entsprechenden Berges relativ zu seiner Umgebung und durch seine Ausmaße. In die relative Lage gehen ein: (1) die Abstände zu den nächsten Tälern und (2) die Anzahl der Berge, die der betreffende Berg dominiert. Hinter dem Konzept der Gewichtung von Knoten steht der Gedanke, daß bestimmte Berge im I-Profil hervorstechen, während andere unbedeutend sind.

Aus dieser Gewichtung lassen sich Parameter ableiten, die aussagen, wie stark die detektierten I-Ereignisse sind. Ein Beispiel ist die Angabe der Kontrastverhältnisse zu beiden Seiten einer Kante. Die Gewichtung kann auch bei der Auswertung komplexer I-Ereignisse verwendet werden.

Für diese Auswertung wurden zwei verschiedene Techniken entwickelt: die Gruppierungen und die Profilverfolgung. Gruppierungen, die auf Gewichtungen aufbauen, können beispielsweise "schwache" Berge mit benachbarten "starken" Bergen gruppieren und aus den betreffenden Gewichten neue Werte für die entstehenden Berg-Gruppen berechnen. Dieser Vorgang kann iterativ fortgesetzt werden, bis keine "schwachen" Berge mehr vorliegen. Bei der Profilverfolgung kann die Gewichtung zur Verknüpfung "gleichstarker" Berge ausgenutzt werden. Derartige Prozesse können als Relaxationsprozesse (vgl. Abschnitt 3.1) formuliert werden.

Im Zusammenhang mit der Profilverfolgung ist ein weiterer Aspekt zu beachten. Durch eine stufenweise Verfolgung von I-Ereignissen, bei der zuerst Knoten nahe der Wurzel von R-Bäumen und anschliessend Knoten, die höher im Baum liegen, verfolgt werden, erfolgt eine hierarchische Verknüpfung. Diese kann als mehrfach angewandte Gruppierung aufgefaßt werden. Zusammenhänge der Gruppierungen auf verschiedenen Ebenen werden durch den S-Baum festgehalten (vgl. Abschnitt 5.3.2). Für die Profilverfolgung wurde ein "Buchhalter"-Modul postuliert, dessen Aufgabe die Überwachung der Zwischenergebnisse ist. Faßt man die Profilverfolgung als eine Gruppierung auf, so kann die Funktion des Buchhalters durch das S-Baum-Modul übernommen werden.

DER INFORMATIONSFLUß ZWISCHEN DEN MODULN

Im Ablauf der Auswertung spielt das S-Baum-Modul eine entscheidende Rolle. Hier fließt alle Information über die Auswertung komplexer I-Ereignisse zusammen. Ein wesentlicher Gedanke des vorliegenden Modells besteht darin, daß die Auswertungs-Moduln ihre Ergebnisse an zwei unterschiedliche Verarbeitungszentren übergeben: an die Skizze und an das S-Baum-Modul.

Das S-Baum-Modul stellt eine Verarbeitungsstelle dar, bei der Entscheidungen nicht aufgrund objektspezifischer Informationen (interne Modelle) getroffen werden, sondern mit Hilfe von Informationen wie: Größe, Form, Komplexität von erzielten I-Ereignissen, sprunghafte Veränderung von Ergebnissen beim Übergang von einer Ebene zur nächsten, usw. Aus den Zusammenhängen zwischen Ergebnis-

sen extrahiert das S-Baum-Modul Tiefenstrukturen, die den Ablauf der weiteren Auswertung entscheiden. Das S-Baum-Modul ist diejenige Instanz, die "weiß", auf welcher Ebene eine Gruppierung arbeitet und wie sie verläuft. Damit hat dieses Modul gewissermaßen den Überblick über die Auswertung und kann z.B. entscheiden, ob eine Gruppierung wiederholt oder abgebrochen werden soll. Diese Entscheidungen können ohne höhere Informationen vorgenommen werden, da aus den komplexen I-Ereignissen hinreichende Informationen für autonome Entscheidungen gewonnen werden. Auf diese Weise kann das S-Baum-Modul bestimmte Gruppierungs-Moduln mehrere Male anstoßen, andere Moduln unterbrechen, bei der Profilverfolgung eine Verknüpfung auf einer feineren Auflösungsebene veranlassen, aber auch A-priori-Information und Ergebnisse höherer Prozesse in den Auswertungsprozess einbeziehen. Durch das S-Baum-Modul wird also die gesamte Auswertung gesteuert, wobei zwei Arbeits-Modi auftreten: (1) das S-Baum-Modul arbeitet autonom, nur aufgrund der ihm zugänglichen Information über die Auswertung oder (2) das Modul wird von höheren Prozessen angesteuert und in seinem Verhalten beeinflusst. Wie alle anderen Moduln übergibt auch das S-Baum-Modul seine Ergebnisse an die Skizze, wo sie als globale Analysen und Deskriptoren der I-Ereignisse vorliegen.

6.2 DIE SKIZZE ALS SEKUNDÄR-REPRÄSENTATION UND IHRE VERARBEITUNG

Aus der Vielzahl der Auswertungs-Ergebnisse von I-Ereignissen läßt sich ein symbolisches Vokabular gewinnen, mit dem eine Hierarchie von Darstellungs-Ebenen für die S-Repräsentation aufgebaut werden kann. Das Konzept der Skizze baut auf der "Erst-Skizze" ('Primal Sketch') auf/44/. Aus der Erst-Skizze werden zwei Gedanken übernommen: (a) der Gebrauch eines symbolischen Vokabulars und (b) die Autonomie der Vorgänge, die die Skizze erstellen. Neben der Erst-Skizze liegen andere Ansätze vor, mit denen das Konzept der Skizze ebenfalls Berührungspunkte hat /7,44/. Unterschiede zu den erwähnten Ansätzen - insbesondere zur Erst-Skizze - bestehen darin, daß (1) der Übergang zur Skizze nicht aus dem Grauwertbild, sondern aus der P-Repräsentation vollzogen wird; (2) daß die

KANTE(1,2,3,4,5,6,7)

- *KONTEXT*=Verweis auf Nachbar-Regionen
- *FORM*= Gerade, Krümmung, ...
- *RICHTUNG*=ϑ
- *STÄRKE*= Rel. Höhe H : H_{mittl}, H_{min}, H_{max}
- *BREITE*= Basisbreite B : B_{mittl}, B_{min}, B_{max}
- *ENDPUNKTE*= Koordinaten des ersten und des letzten K-Elementes
- *TYP*= je nach Art der K-Elemente (konvex, konkav, ...)

LINIE (1,2,3,4,5,6,7,8)

- *KONTEXT*= Verweis auf Nachbar-Regionen
- *FORM*= Gerade, Krümmung , ...
- *RICHTUNG*=ϑ
- *STÄRKE*= Rel. Höhe H : H_{mittl}, H_{min}, H_{max}
- *BREITE*= Basisbreite B : B_{mittl}, B_{min}, B_{max}
- *ENDPUNKTE*= Koordinaten der Endpunkte
- *RECHTE FLANKE*= Typ der rechten Flanke (Steigung , Länge ,...)
- *LINKE FLANKE*= Typ der linken Flanke (Steigung , Länge , ...)

REGION(1,2,3,4,5,6)

- *KONTEXT*= Verweis auf Nachbar-Regionen.
- *TEXTUR*= Textur-Deskriptoren (Merkmalsvektoren)
- *FORM* = Form-Deskriptoren
- *FLÄCHE*= Summe der Basisbreiten
- *KONTUR*= Radiusfunktion um Schwerpunkt zur Kontur
- *SCHWERPUNKT*= ausgezeichnete Koordinaten zur örtl. Festlegung

ABBILDUNG 51 : ELEMENTE EINES VOKABULARS FÜR DIE SKIZZE
KANTE(1,2,...,7) SETZT SICH AUS K-ELEMENTEN ZUSAMMEN (VGL. ABB. 29)
LINIE(1,2,...,8) SETZT SICH AUS L-ELEMENTEN ZUSAMMEN

Skizze aus mehreren Schichten besteht und (3) daß das verwendete Voka - bular reichhaltiger ist.

DIE SPEICHERUNG DER ERGEBNISSE

In der Skizze werden alle Auswertungs-Ergebnisse gespeichert. Dabei sind einige Forderungen an die Skizze zu stellen:

- Alle I-Ereignisse müssen topologisch richtig gespeichert werden.
- Die I-Ereignisse werden durch ein symbolisches Vokabular darge - stellt, das auch Merkmale der Ereignisse enthält. Ein Beispiel für

ein Element eines solchen Vokabulars ist in Abb. 29 enthalten. Abbildung 51 zeigt weitere Beispiele dafür, wie ein solches Vokabular aussehen könnte.

- Die I-Ereignisse müssen in Schichten nach Typen geordnet übereinanderliegen, so daß Informationen zwischen den Schichten hin und her fließen können. Die Daten müssen sich gegenseitig beeinflussen, so daß sich als Folge der Wechselwirkung der Daten ein konsistentes Ergebnis einstellt.

- Die Verarbeitung innerhalb der Skizze muß (1) zu einer Bereinigung der Ergebnisse und (2) zu einer Segmentation und Merkmalsextraktion führen.

- Die Skizze dient als Schnittstelle zu höheren Prozessen, die ausschließlich mit den Daten der Skizze arbeiten und mit Hilfe des Vokabulars eine Analyse der segmentierten Regionen vornehmen. Rückwirkungen höherer Prozesse auf die Erzeugung der Skizze erfolgen nur, wenn mehr Information für die Analyse benötigt wird. In diesem Fall kann über das S-Baum-Modul eine weitere Auswertung veranlaßt werden.

Wie sieht nun die Skizze aus? Sie besteht aus drei Schichten und einer Vielzahl von Säulen, die jeweils einen Ort im Grauwertbild repräsentieren. Ein Ort kann ein einzelner Bildpunkt sein; aus Aufwandsgründen erscheint es günstiger, jeweils mehrere benachbarte Bildpunkte als einen Ort aufzufassen, da auch einfache I-Ereignisse stets über mehreren Bildpunkten liegen. In den drei Schichten werden die einfachen, komplexen und hyperkomplexen I-Ereignisse durch ein Vokabular dargestellt. In jeder der Schichten kann für einen Ort mehr als ein Ergebnis vorliegen. Mehrfache Ergebnisse werden in den Säulen gespeichert. Die Entwicklung einer konkreten Datenstruktur für eine Implementierung wird hier nicht angestrebt, da es zunächst darauf ankommt, Grundzüge eines geeigneten Modelles herauszuarbeiten. Ein wesentlicher Bestandteil des Modells ist die Vorstellung, daß sich in der Skizze stets nur soviel Information befindet, wie für die Analyse benötigt wird. Durch den schichtenweisen Aufbau der Skizze ist es möglich, zu Beginn der Szenenanalyse

nur Auswertungen von Grobstrukturen in die Skizze einzutragen. Mit Hilfe dieser Auswertungs-Ergebnisse werden Regionen ausgewählt, die durch eine feinere Auswertung erfaßt werden. Auf diese Weise können sowohl Speicheraufwand als auch Verarbeitungsaufwand verringert werden, ohne auf hohe Auflösung und sorgfältige Analysen verzichten zu müssen.

DIE BEREINIGUNG DER SKIZZE

I-Ereignisse, die in einer Säule untereinander liegen, kennzeichnen die verschiedenen Ergebnisse für den entsprechenden Ort und können sich gegenseitig widersprechen. Dagegen ist zu erwarten, daß Ergebnisse in Nachbarsäulen redundant sind und sich gegenseitig unterstützen. An dieser Stelle kann nun das Konzept der Kooperation und Konkurrenz / 2 / integriert werden, d.h. die verschiedenen Auswertungs-Ergebnisse wetteifern gegeneinander oder unterstützen sich gegenseitig, bis ein konsistentes Ergebnis vorliegt. Kooperation und Konkurrenz werden formalisiert durch Relaxatationsprozesse /62/, die auf der Skizze in zwei Richtungen arbeiten, indem Kooperation zwischen den Säulen und Konkurrenz innerhalb von Säulen stattfindet.

Die oben eingeführte Gewichtung kann dazu benützt werden, Relaxationsprozesse zu entwickeln. Hierfür müssen für die verschiedenen I-Ergebnisse Kompatibilitätsmatrizen aufgestellt werden. Da nicht allzuviele Typen von I-Ereignissen für die Skizze benötigt werden, sind diese Matrizen klein. Mit Hilfe der Relaxation wird die Skizze solange bereinigt, bis sich ein konsistentes Ergebnis eingestellt hat.

DIE AUSWAHL EINER BESCHREIBUNGSEBENE

Nach Bereinigung der Skizze liegt eine symbolische Beschreibung auf mehreren Ebenen vor, in denen Konturelemente, Texturelemente, Konturen und Regionen zusammen mit ihren Merkmalen enthalten sind. Aus dieser hierarchischen Darstellung muß noch eine geeignete Beschreibungsebene ausgewählt werden, die die Information für die Analyse durch höhere Prozesse liefert. Die Hierarchie von I-Ereignissen erlaubt, an dieser

Stelle das Konzept des Spaltens und Schmelzens einzubeziehen /34/. In der Skizze liegen in den Schichten Symbole vor, die nach unten hin immer größeren I-Ereignissen entsprechen. Diese I-Ereignisse können nun nach oben und unten durch Spalten und Schmelzen zu einer Beschreibung zusammengesetzt werden, der durch höhere Prozesse ein internes Modell zugeordnet werden kann. So ist es z.B. möglich, eine Grobstruktur zusammen mit Details, die an ausgesuchten Punkten der Grobstruktur liegen, auszuwählen und an die höheren Prozesse weiterzugeben. Ebenso kann die Analyse auf der Ebene hyperkomplexer oder komplexer Ereignisse beginnen und interne Modelle zuordnen, die dann auf der Ebene der einfachen Ereignisse verifiziert werden. Mit der Zuordnung zu internen Modellen gewinnt man eine Terminal-Repräsentation. Durch die oben beschriebene Arbeitsweise des Modells ist es möglich, eine T-Repräsentation aufzubauen, deren Auflösung örtlich verschieden ist: detaillierte Beschreibungen treten in den Zonen des Interesses auf, während weniger wichtige Regionen nur als Grobstrukturen vorliegen.

6.3 DISKUSSION

Mit den Ausführungen dieses Kapitels liegt ein gedankliches Modell vor, das die Arbeitsweise eines autonomen, verteilten Systems niederer Prozesse beschreibt. Viele der hier eingeführten Begriffe bedürfen noch einer Präzisierung durch Formalismen und Algorithmen. Das Modell bietet einen Rahmen für fortführende Arbeiten.

Wodurch ist das Modell gekennzeichnet? Es gliedert den Vorgang der Szenenanalyse in mehrere Zwischenschritte auf. Wesentlich sind:

1) Der <u>Übergang vom Grauwertbild zur P-Repräsentation</u>.

   Hierbei sind zu beachten:

   - der Aufbau einer hierarchischen Darstellung;
   - das Sammeln von Kontext-Information;
   - die Erhaltung der Struktur der Grauwertverteilung.

   Dieser Übergang wird durch die Konstruktion von Relational-Bäumen

vollzogen, die eine Beschreibung der Intensitätsprofile von Bildzeilen liefern.

2) Der Übergang von der P-Repräsentation zur S-Repräsentation.

Hierbei sind zu beachten:

- die Zerlegung der P-Repräsentation in Teilstrukturen, die Intensitäts-Ereignissen entsprechen;
- die Beschreibung dieser Ereignisse durch ein reichhaltiges symbolisches Vokabular;
- die Auswertung der P-Repräsentation nach verschiedenen Kriterien;
- die Erfassung von Zusammenhängen von Ergebnissen durch das S-Baum-Modul;
- die Erfassung von Ergebnissen in der Skizze.

3) Die Bereinigung der S-Repräsentation.

Hierbei sind zu beachten:

- für einen Ort liegen mehrere Auswertungs-Ergebnisse vor;
- endgültige Ergebnisse werden durch Kooperation und Konkurrenz der Daten gewonnen;
- die Schichtenstruktur dieser Repräsentation erlaubt wechselnde Beschreibungsebenen.

4) Der Übergang von der S-Repräsentation zur T-Repräsentation.

Hierbei sind zu beachten:

- das reichhaltige Vokabular und die Beschreibungsebenen der S-Repräsentation unterstützen die Analyse;
- die Darstellung der T-Repräsentation kann an verschiedenen Orten verschiedene Auflösung haben - je nach Wahl der Beschreibungsebene;
- Für jeden Ort liegt nur noch eine Beschreibung in Form einer

Zuordnung zu einem internen Modell vor.

Ein solches Modell erfordert einen hohen Aufwand - insbesondere einen hohen Speicheraufwand. Die Rechtfertigung eines solchen Aufwandes ergibt sich aus der Tatsache, daß einfache Verfahren für die Analyse natürlicher Szenen nicht ausreichen /72/. Man hat - insbesondere in den ersten Jahren der Verfahrensentwicklung - versucht, mit einigen wenigen Operationen brauchbare Ergebnisse zu erzielen. Heute weiß man, daß die Probleme nicht gelöst werden, indem man immer bessere Operatoren entwickelt (etwa zur Kantendetektion). Statt dessen muß:

- die Information, die einzelne Verfahren liefert, besser ausgeschöpft werden;
- eine Kombination von Verfahren eingesetzt werden, die sich gegenseitig ergänzen und unterstützen.

Als Folge der ersten Forderung wird ein symbolisches Vokabular notwendig, dessen Elemente mehr Information enthalten als nur eine Zahl, die einem Operatorwert entspricht. Elemente eines solchen Vokabulars erfassen die tatsächliche Struktur von I-Ereignissen und erleichtern damit die Verknüpfung der Elemente untereinander. Eine Vereinfachung des hier vorgeschlagenen Vokabulars ist durch Weglassen von Information leicht zu vollziehen. Hier kam es im Gegenteil darauf an, zu zeigen, wie mehr Informationen über I-Ereignisse gewonnen und dargestellt werden kann.

Die zweite Forderung setzt Mechanismen voraus, die ein sinnvolles Zusammenspiel mehrerer Verfahren erlaubt. Hierfür wurde ein Konzept entwickelt, das darauf beruht, die Ergebnisse der Auswertung an zwei Stellen zu erfassen: im S-Baum-Modul und in der Skizze. Während für die Skizze die Ergebnisse selbst interessant sind, kommt es im S-Baum-Modul darauf an, wie sich die Ergebnisse entwickeln. Durch diesen dynamischen Aspekt ist es möglich, über den Zusammenhang zwischen den Ergebnissen die Auswertungs-Moduln zu kontrollieren. Treten z.B. im Ablauf einer Gruppierung sprungartige Veränderungen der Ergebnisse auf, so läßt sich vermuten, daß die Gruppierung eine falsche Region erfaßt und verknüpft hat und deswegen abgebrochen werden muß. Auf diese Weise sorgt das S-

Baum-Modul dafür, daß die Skizze nicht mit einem Übermaß an Ergebnissen vollgeschrieben wird. Auch hier kann durch Weglassen von Moduln eine Vereinfachung erzielt werden.

Das vorliegende Modell weist Ähnlichkeiten mit anderen Ansätzen auf /32,44/, verwendet jedoch unterschiedliche Techniken. Das Modell niederer Prozesse von MARR /44/ vollzieht direkt den Übergang vom Grauwertbild zur "Erst-Skizze", d.h. als P-Repräsentation dient das Grauwertbild selbst. Um zu einem symbolischen Vokabular zu gelangen, wird das Bild mit verschiedenen Operatoren gefaltet und für jeden Bildpunkt aus den Ergebnissen der Faltungen (bis zu 32) das beste ausgesucht. Neben dem hohen Aufwand dieses Ansatzes liegt ein Problem in der Auswahl eines Ergebnisses aus einer Reihe von Faltungen, da nicht sicher ist, daß durch die getroffene Wahl die tatsächliche Struktur der Grauwertverteilung ermittelt wurde.

Das vorliegende Modell zeigt den Zusammenhang zwischen den Techniken, die in dieser Arbeit entwickelt wurden. Durch seinen Aufbau erklärt das Modell:

- wie es möglich ist, Segmentation und Merkmalsextraktion auf verschiedenen Ebenen zu vollziehen, ohne alle Ebenen unterhalb der gewünschten abzuarbeiten;
- wie globale und lokale Analysen im Wechselspiel ablaufen können;
- wie niedere Prozesse autonom ohne Einfluß höherer Prozesse zu einer Segmentation und Merkmalsextraktion eines Grauwertbildes gelangen können.

7. ZUSAMMENFASSUNG

Thema der vorliegenden Arbeit sind Probleme der Szenenanalyse. Hautpziel ist die Entwicklung leistungsfähiger autonomer niederer Prozesse, die ein Grauwertbild einer natürlichen Szene bis zur Segmentation und Merkmalsextraktion verarbeiten, ohne auf höhere Prozesse zurückgreifen zu müssen.

Mechanismen der Szenenanalyse beim Menschen (Kapitel 2.) zeigen wichtige Phänomene. Die Verarbeitung visueller Reize läuft beim Menschen in mehreren Hirnregionen ab, die z.T. hintereinander, z.T. nebeneinander geschaltet sind. Wesentlich ist, daß im Cortex eine Analyse der Form vollzogen wird, während im Colliculus Lage, Größe und zeitliche Veränderung ermittelt werden. Die Topologie der Retina bleibt dabei stets erhalten. Bei der Verarbeitung der Reize beeinflussen sich benachbarte Zellen gegenseitig durch Erregung und Hemmung. Die Trennung von Hintergrund und Figur ist unabhängig vom Formensehen und kann auf verschiedenen Auflösungsebenen erfolgen.

Bei der Szenenanalyse im Computer (Kapitel 3.) liegen für die Segmentation zwei grundsätzlich verschiedene Verfahren vor: man kann entweder alle Konturlinien ermitteln oder Regionen durch ihre Flächeneigenschaften (z.B. Grauwerte oder Textur) ermitteln. Im ersten Fall erfolgt eine Verarbeitung über Kanten- und Liniendetektion, Linienverfolgung und Markierung von Linien, im zweiten Fall werden Regionen, die sich aus einer Initialsegmentation ergeben, durch Spalten und/oder Verschmelzen verarbeitet, bis die gewünschte Segmentation erreicht ist. Für den zweiten Ansatz ist die Erfassung von Texturen ein zentrales Problem. Als Texturelemente werden üblicherweise Kantenelemente, Lauflängen, Grauwertpaare oder lokale Extremwerte verwendet. Um Entscheidungen nicht allein aufgrund lokaler Informationen treffen zu müssen, werden hierarchische Repräsentationen eingesetzt, die die Auswertung durch die Bereitstellung globaler Information unterstützen. Als Datenstrukturen werden hierfür Pyramidensysteme oder Bäume eingesetzt. In beiden Fällen ergibt sich die hierarchische Repräsentation durch eine Zerlegung der Bilder aufgrund einer fest vorgegebenen Geometrie.

Zur Erprobung einiger Verfahren wurden Experimente durchgeführt zur: (1) Kompression im Ortsbereich, (2) Kantendetektion mit Operatoren der Literatur und einem eigenen Kantendetektor und (3) Segmentation durch Histogrammanalysen. Aus den Ergebnissen der Experimente ergibt sich eine Kritik der untersuchten Verfahren. Um eine Beurteilung verschiedener Verfahren von einem einheitlichen Standpunkt aus zu ermöglichen, wird der Vorgang der Szenenanalyse als eine Folge von Übergängen von Repräsentationen postuliert. Als wichtige Typen von Repräsentationen werden die Primär- , Sekundär- und die Terminal-Repräsentation eingeführt, die alle unterschiedliche Aufgaben haben. Für die Übergänge zwischen diesen Repräsentationen werden Verarbeitungsprinzipien in einem Katalog zusammengestellt.

Unter Berücksichtigung dieser Prinzipien wird ein neuartiges Konzept (Kapitel 4.) einer symbolischen Repräsentation von Grauwertbildern entwickelt. Die Grundidee besteht darin, in einer Primär-Repräsentation die Struktur der Grauwertverteilung eines Bildes in hierarchischer Weise zu beschreiben, um so Kontext-Information für die nachfolgende Auswertung bereitzustellen. Als konkrete Datenstruktur wird hierfür eine Baumstruktur, Relational-Baum genannt, entwickelt. Relational-Bäume liefern eine Beschreibung von Intensitätsprofilen von Bildzeilen, wobei die Struktur der Profile weitgehend erhalten bleibt. Die Baumstruktur wird erzeugt, indem Bildpunkte nach kontext-abhängigen Kriterien zusammengefaßt werden. Zur Konstruktion der Baumstrukturen wird ein Kelleralgorithmus entwickelt, der in einem Durchlauf die gewünschte Beschreibung liefert.

Die Auswertung der Relational-Bäume (Kapitel 5.) erfolgt nach verschiedenen Gesichtspunkten. Zunächst wird anhand von Rekonstruktionen der Intensitätsprofile aus den Relational-Bäumen gezeigt, in welchem Umfang die Struktur der Profile erhalten bleibt. Anschließend werden Strukturen von Elementen der Grauwertbilder (Kanten-, Linien-, Textur-Elemente) untersucht und gezeigt, wie diese aus Relational-Bäumen ermittelt und beschrieben werden können. Die Aufgabe der Flächenbildung wird durch Gruppierungen der Elemente der Relational-Bäume vorgenommen. Die Ermitt-

lung der Zusammenhänge der Ergebnisse, die sich durch mehrfache Gruppierungen ergeben, liefert Tiefenstrukturen. Für die Texturanalysen wird ein rekursives Texturmodell mit einer Hierarchie von sich einschließenden Texturelementen zugrundegelegt und anhand von Experimenten gezeigt, wie solche Texturen beschrieben werden können. Um die eindimensionale Verarbeitung in Zeilenrichtung zu ergänzen, wird eine Verarbeitung quer zu den Zeilen eingeführt (Profilverfolgung). Hier werden Teilstrukturen über mehrere Zeilen hinweg verfolgt und miteinander verknüpft. Für alle diese Auswertungstechniken werden Algorithmen entwikkelt und ihre Arbeitsweise durch Experimente demonstriert.

In einem gedanklichen Modell (Kapitel 6.) wird der Zusammenhang zwischen den verschiedenen oben entwickelten Techniken hergestellt, dabei werden auch einige Verfahren der Literatur in das Modell mit einbezogen. Das Modell greift noch einmal die Darstellung der Szenenanalyse als eine Folge von Übergängen von Repräsentationen auf. Der Übergang vom Grauwertbild zur Primär-Repräsentation (im Modell: Relational-Bäume) stellt eine Aufbereitungsphase dar, bei der die Struktur der Grauwertverteilung erhalten bleibt. Die Primär-Repräsentation wird beim Übergang zur Sekundär-Repräsentation (im Modell: Skizze) in eine Vielzahl von Teilstrukturen zerlegt, die Intensitäts-Ereignisse entsprechen und durch ein reichhaltiges symbolisches Vokabular beschrieben werden. Durch Analyse von Tiefenstrukturen wird der Vorgang der Auswertung gelenkt und verhindert, daß zuviele Ergebnisse in die Skizze eingetragen werden. Durch die Organisation der Skizze können für eine Ort im Bild mehrere, auch widersprüchliche Auswertungs-Ergebnisse vorliegen, die in einer anschließenden Bereinigungsphase entfernt werden. Höhere Prozesse, die eine Analyse vornehmen, greifen auf die Sekundär-Repräsentation zu und ordnen Segmenten der Skizze interne Modelle zu.

8. LITERATURVERZEICHNIS

/1/ ARBIB, M.A. — The Metaphorical Brain. J. Wiley & Sons, Inc., N.Y. 1972

/2/ ARBIB, M.A. — Segmentation, Slides, And Cooperative Computation. COINS Tech. Report, Univ. of Mass., Amherst, 1975

/3/ ARBIB, M.A. — Artifical Intelligence And Brain Theory: Unities And Diversities. COINS Tech. Report 75c-8, Univ. of Mass., Amherst, 1975

/4/ ARBIB, M.A. RISEMAN,E.M. — Computational Techniques In Visual Systems,Part I.:The Overall Design. COINS Tech. Report 76-10, Univ. of Mass., Amherst, July 1976

/5/ BAJCSY, R. — Computer Description Of Textured Surfaces. Proc. of 3rd IJCAI, August 1973

/6/ BARROW, H.G. POPPLESTONE, R.J. — Relational Descriptions In Picture Processing. in: MELTZER, B. & MICHIE, D. (Hrsg.) Machine Intelligence 6, American Elesevier, N.Y., 1971, S. 377 - 396

/7/ BERTELSMEIER, R. RADIG, B. — Kontextunterstützte Analyse von Szenen mit bewegten Objekten. in: NAGEL, H.H. (Hrsg.): Digitale Bildverarbeitung GI/NTG Fachtagung, München, März 1977 Informatik Fachberichte 8, Springer Verlag, S. 101 - 128

/8/ BRICE, C.R. FENNEMA, C.L. — Scene Analysis Using Regions. Artifical Intelligence Group, Tech. Note 17, Stanford Research Institute, Stanford, 1970

/9/ BULLOCK, B. — The Performance Of Edge Operators On Images With Texture. Hughes Res. Lab., Tech. Report, Malibu, Oct. 1974

/10/ CARLTON, S.G. MITCHELL, O.R. — Image Segmentation Using Texture And Gray Level. Proc. of Conf. on Pattern Rec. and Image Processing, Troy, June, 1977, S. 387 - 391

/11/ CHOW, C.K.
KANEKO, T.
Automatic Boundary Detection Of The Left Ventricle From Cineangiograms.
Comput. Biomed. Res. 5 (1972), S. 388 - 410

/12/ CLOWES, M. B.
On Seeing Things.
Artifical Intelligence 2 (1971) 1, S. 79 - 116

/13/ DAVIS, L.S.
A Survey Of Edge Detection Techniques.
Computer Graphics & Image Processing 4(1975), S. 248 - 270

/14/ DEUTSCH, E.S.
BELKNAP, N.J.
Texture Description Using Neighborhood Information.
Computer Graphics & Image Processing 1(1972), S. 145 - 168

/15/ DIDDAY, R.S.
ARBIB, M.A.
Eye Movements And Visual Perception: A "Two Visual System" Model.
COINS Tech. Report, Univ. of Mass., Amherst, 1973

/16/ DUDA, R.O.
Pattern Classification And Scene Analysis.
J. Wiley & Sons, N.Y., 1973

/17/ EHRICH, R.W.
FOITH, J.P.
Representation Of Random Waveforms By Relational Trees.
IEEE Trans. Computers C - 25 (1976) July, S. 725 - 736

/18/ EHRICH, R.W.
FOITH, J.P.
A View Of Texture Topology And Texture Description.
Computer Graphics & Image Processing (1978), to appear

/19/ EHRICH, R.W.
FOITH, J.P.
Topology And Semantics Of Intensity Arrays.
in /58/

/20/ EHRICH, R.W.
Detection Of Global Edges In Textured Images.
ECE Tech. Report, Univ. of Mass., Amherst, Nov. 1975

/21/ EHRICH, R.W.
A Symmetric Hysteresis Smoothing Algorithm That Preserves Principal Features.
CS Tech. Report, Virginia Polytechnic Institute and State Univ., Blacksburg, 1977

/22/ FOITH, J.P. Entwicklung von Modellvorstellungen des menschlichen Erkennungssystems.
IITB-Bericht, Fraunhofer-Gesellschaft, Karlsruhe, 1973

/23/ FOITH, J.P. Summary On Differentiating And Image Compressing Operators.
ECE Tech. Report, Univ. of Mass., Amherst, April 1975

/24/ FREUDER, E.C. A Computer System For Visual Recognition Using Active Knowledge.
MIT AI Lab., Tech. Report TR-345, Cambridge, 1976

/25/ GALLOWAY, M. Texture Analysis Using Gray Level Run Lengths.
Computer Graphics and Image Processing 4 (1975), S. 172 - 179

/26/ GRAHAM, C.H. Vision And Visual Perception.
John Wiley & Sons, N.Y., 1965

/27/ GREGORY, R.L. Auge und Gehirn.
Zur Psychologie des Sehens.
Fischer Verlag 1972

/28/ GUZMAN, A. Computer Recognition Of Three-Dimensional Objects In A Visual Scene.
MIT Thesis, MIT AI Lab., Cambridge, 1968

/29/ HANSON, A.R., RISEMAN, E.M. Preprocessing Cones: A Computational Structure For Scene Analysis.
COINS Tech. Report, Univ. of Mass., Amherst 1974

/30/ HANSON, A.R. RISEMAN, E.M. The Design Of A Semantically Directed Vision Processor.
COINS Tech. Report 75c-1, Univ. of Mass., Amherst 1975

/31/ HANSON, A.R. RISEMAN, E.M. NAGIN, P. Region Growing In Textured Outdoor Scenes.
COINS Tech. Report 75c-2, Univ. of Mass., Amherst 1975

/32/ HANSON, A.R. RISEMAN, E.M. Representation And Control In The Construction Of Visual Models.
COINS Tech. Report 76-9, Univ. of Mass., Amherst, 1976

/33/ HARALICK, R.M.
SHANMUGAN, K.
DINSTEIN, I.
Textural Features For Image Classification.
IEEE Trans. SMC. 3 (1973), S. 610 - 621

/34/ HOROWITZ, S.L.
PAVLIDIS, T.
Picture Segmentation By A Directed Split-and Merge Procedure.
Proc. 2 IJCPR, Copenhagen, 1974, S. 497 - 498

/35/ HORN, B.K.P.
Obtaining Shape From Shading Information.
in /71/

/36/ HORN, B.K.P.
Image Intensity Understanding.
AI Memo 335, MIT AI Lab., Cambridge, Aug. 1975

/37/ HUECKEL, M.
An Operator Which Locates Edges In Digitized Pictures.
Journal of the ACM 18 (1971) 1, S. 113 - 125

/38/ JULESZ, B.
Foundations Of Cyclopean Perception.
Univ. of Chicago Press, Chicago, 1971

/39/ KELLY, M.D.
Edge Detection In Pictures By Computer Using Planning.
Machine Intelligence 6 (1971), S. 379 - 409

/40/ KLINGER, A.,
DYER, C.R.
Experiments On Picture Representation Using Regular Decomposition.
Tech. Report UCLA-Eng 7494, Univ. of Cal., Los Angeles, 1974

/41/ KRAKAUER, L.J.
Computer Analysis Of Visual Properties Of Curved Objects.
MIT TR-82, Project MAC, MIT AI Lab., Cambridge, May 1971

/42/ MONTANARI, U.
On The Optimal Detection Of Curves In Noisy Pictures.
Journal of the ACM 14 (1971) 5

/43/ MARR, D.
Analyzing Natural Images.
MIT AI Memo 334, MIT AI Lab., Cambridge, 1975

/44/ MARR, D.
Early Processing Of Visual Information.
MIT AI Memo 340, MIT AI Lab., Cambridge, 1975

/45/ MARR, D. Representing Visual Information.
in /58/

/46/ MARTELLI, A. Edge Detection Using Heuristic Methods.
Computer Graphics & Image Processing 1 (1972) 2, S. 169 - 182

/47/ MINSKY, M. A Framework For Representing Knowledge.
MIT AI Memo 306, MIT AI Lab., Cambridge, 1974

/48/ MITCHEL, O.R.
MYERS, C.R.
BOYNE, W. A Max-Min Measure For Image Texture Analysis.
IEEE Trans. Comp. C-26, 1977, S. 408 - 414

/49/ MUERLE, J.L.
ALLEN, D.C. Experimental Evaluation Of Techniques For Automatic Segmentation Of Objects In a Complex Scene.
In: CHENG, G.C. et al. (Hrsg.), Pictorial Pattern Recognition, Thomson, Washington, 1968, S. 3 - 13

/50/ NIEMANN, H. Methoden der Mustererkennung.
Akademische Verlagsgesellschaft, Frankfurt/M. 1974

/51/ NOTON, D.
STARK, L. Scanpaths In Saccadic Eye Movements While Viewing And Recognizing Patterns.
Vision Research 11 (1970), S. 929 - 941

/52/ NORMAN, D.A. Models For Human Memory.
In: Encylopedia of Linquistics, Information and Control. Pergamon Press, Oxford 1967, S. 311 - 313

/53/ OHLANDER, R.B. Analysis Of Natural Scenes.
Dept. of Comp. Sc., Carnegie-Mellon-University, Pittsburgh, Dissertation, 1975

/54/ OLSON, D. Cognitive Development.
Academic Press, 1970

/55/ Paul, D. Das visuelle System des Menschen gesehen unter technischen Aspekten.
IITB-Bericht, Fraunhofer-Gesellschaft, Karlsruhe, 1975

/56/ PAVLIDIS, T.
Segmentation Of Pictures And Maps Through Functional Approximation.
Computer Graphics and Image Proc. 1 (1972), S. 360 - 372

/57/ RISEMAN, E.M.
ARBIB, M.A.
Computational Techniques In Visual Systems, Part II.: Segmenting Static Scenes.
COINS Tech. Report 76-11, Univ. of Mass., Amherst, July 1976

/58/ RISEMAN, E.M.
HANSON, A.R.
(Hrsg.)
Proceedings Of The Workshop On Computer Vision Systems.
held at Univ. of Mass., Amherst, June 1-3, 1977

/59/ ROSENFELD, A.
THOMAS, R.,
LEE, Y.
Edge And Curve Enhancement In Digital Pictures.
Picture Processing Lab., Tech. Report, Univ. of Maryland, College Park, 1969

/60/ ROSENFELD, A.
LEE, Y.
Edge And Curve Detection, Further Experiments.
IEEE Trans. Computers C-21, 1972, S. 677 - 715

/61/ ROSENFELD, A.
HUMMEL, R.A.
ZUCKER, S.W.
Scene Labelling By Relaxation Operations.
Picture Processing Lab., Tech. Report 379, University of Maryland, College Park, 1975

/62/ ROSENFELD, A.
Iterative Methods In Image Analysis.
Proc. of Conf. on Pattern Rec. and Image Processing, Troy, June, 1977, S. 14 - 18

/63/ SCHACHTER, B.J.
LEV, A.
ZUCKER, S.W.
ROSENFELD, A.
An Application Of Relaxations Methods To Edge Reinforcement.
CS Tech. Report 476, Univ. of Maryland, College Park, Aug. 1976

/64/ v. SEELEN, W.
Zur Informationsverarbeitung im visuellen System der Wirbeltiere.
Kybernetik 7 (1970), S. 43 - 60

/65/ SHIRAI, J.
Analyzing Intensity Arrays Using Knowledge About Scenes.
in /71/

/66/ TANIMOTO, S.
PAVLIDIS, T.
A Hierarchical Data Structure For Picture Processing.
Computer Graphics and Image Processing 4 (1975), S. 104 - 119

/67/ TSUJI, S.
TOMITA, F.
A Structural Analyzer For A Class of Texture.
Computer Graphics and Image Processing 2 (1973), S. 216 - 331

/68/ UHR, L.
Layered Recognition Cone Networks That Preprocess, Classify, And Describe.
IEEE Trans. Computers C-21 (1972), S. 758 - 768

/69/ WALTZ, D.
Understanding Line Drawings Of Scenes With Shadows.
in /71/

/70/ WERBLIN, F.S.
The Control Of Sensitivity In The Retina.
Scientific American (1973) 1, S. 71 - 79

/71/ WINSTON, P.H. (Hrsg.)
The Psychology Of Computer Vision.
McGraw-Hill Book Company, N.Y., 1975

/72/ WINSTON, P.H.
Artificial Intelligence.
Addison-Wesley Publ. Comp., Reading, Mass., 1977

/73/ YARBUS, A.L.
Eye Movements And Vision.
Plenum Press, 1967

/74/ TENENBAUM, J.M.
BARROW, H.G.
Experiments In Interpretation-Guided Segmentation.
Tech. Note 123, AI Center, Stanford Res. Inst., 1976

/75/ YAKIMOVSKY, Y.
FELDMAN, J.
A Semantics-Based Decision Theory Region Analyzer.
Proc. 3rd Int. Joint Conf. on Artificial Intelligence, 1973, S. 580 - 588

/76/ ZINCHENKO, V.P.
VERGILES, N.Y.
Formation Of Visual Images.
Translated From Russian, Consultant Bureau, N.Y., 1972

/77/ ZUCKER, S.W.
ROSENFELD, A.
DAVIS, L.S.
Picture Segmentation By Texture Discrimination.
Picture Processing Lab. Tech. Report-356 Univ. of Maryland, College Park, 1975

/78/ ZUCKER, S.W.
HUMMEL, R.A.
ROSENFELD, A.
Applications Of Relaxation Labelling, Line And Curve Enhancement.
CS Tech. Report-419, Univ. of Maryland, College Park, Oct. 1975

/79/ ZUCKER, S.W.
KRISHNAMURTHY, E.
HAAR, R.L.
Relaxation Processes For Scene Labelling;Convergence, Speed, And Stability.
CS Tech. Report-477, Univ. of Maryland, College Park, Aug. 1976

/80/ ZUSNE, L.
Visual Perception Of Form.
Academic Press, N.Y., 1970

SCHLUSSBEMERKUNG

Die vorliegende Arbeit wurde am Institut für Informationsverarbeitung in Technik und Biologie (IITB) der Fraunhofer-Gesellschaft in Karlsruhe angefertigt. Wesentliche Teile der Arbeit entstanden während zweier Studienaufenthalte in den Vereinigten Staaten von Amerika : am Department of Electrical and Computer Engineering, University of Massachusetts in Amherst, Ma. und am Department of Computer Science, Virginia Polytechnic Institute and State University in Blacksburg, Va. Für die Ermöglichung dieser Aufenthalte danke ich Herrn Dr. A. Schief, dem Direktor des IITB. Der Fraunhofer-Gesellschaft und der National Science Foundation danke ich für die teilweise Finanzierung.

Bei beiden Studienaufenthalten war Prof. Dr. R.W. Ehrich mein Gastgeber. Er sorgte für ein anregendes, fruchtbares Arbeitsklima und stellte großzügig die Arbeitsmittel bereit. Von ihm gingen wesentliche Fragestellungen aus, und er hat mit einer Vielzahl von Anregungen zum Gelingen dieser Arbeit beigetragen. Dafür gilt ihm mein herzlichster Dank.

Herrn Prof. Dr. H. Niemann und Herrn Prof. Dr. H.J. Schneider danke ich für die Übernahme der Referate. Herr Prof. Dr. H. Niemann hat durch seine konstruktive Kritik zur Verbesserung von Inhalt und Form der Arbeit beigetragen, wofür ich ihm herzlich danke.

Dem IITB danke ich für die großzügige Unterstützung bei der Anfertigung dieser Arbeit im Rahmen meiner Tätigkeit als wissenschaftlicher Mitarbeiter. Sein Direktor, Herr Dr. A. Schief, hat in dankenswerter Weise meine Arbeit gefördert. Meinen Institutskollegen - allen voran Herrn Dr. D. Paul und Herrn Dr. G. Zimmermann - danke ich für die vielen anregenden Diskussionen über das Thema dieser Arbeit.

LEBENSLAUF

23. 9. 1945	Geburt in Eugendorf/Kreis Salzburg/Österreich als Sohn des Lehrers Hermann Foith und der Gertrud Foith, geb. Weimer
1949	Umzug der Familie nach Karlsruhe
1952 - 1956	Besuch der Volksschule in Karlsruhe
1956 - 1965	Besuch des Goethe-Gymnasiums in Karlsruhe
1965 - 1967	Militärdienst in einer Fernmelde-Einheit Abgang als Leutnant der Reserve
1967 - 1969	Studium an der Universität Karlsruhe Fachrichtung : Mathematik Abschluß : Vordiplom
197o - 1972	Studium an der Universität Karlsruhe Fachrichtung : Informatik Abschluß : Diplom
1972 - heute	Wissenschaftlicher Mitarbeiter am Institut für Informationsverarbeitung in Technik und Biologie (IITB) der Fraunhofer-Gesellschaft e.V., Karlsruhe
Jan. - Juli 1975	Studienaufenthalt am Dept. of Electrical and Computer Engineering, University of Massachusetts, Amherst, Ma.
Nov. 76 - Jan. 77	Studienaufenthalt am Dept. of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, Va.

Offsetdruck-Fotodruck
J. Hogl, 8520 Erlangen, Hauptstr. 109 (Nähe Martin-Luther-Platz) Tel. 22990